

Tue Sep 21 09:16:38 2004

us-10-009-945-4.rspt

Page 1

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 21, 2004, 07:36:41; Search time 91.0211 Seconds
(without alignments)
2592.890 Million cells updates/sec

Title: US-10-009-945-4
Perfect score: 4038
Sequence: 1 MSNPRRRNGPVKRLTVLC.....EKLYEKLLTAIEETGFAVE 748

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvtrius:*
16: sp_bacteriap:*
17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	3041	75.3	728	11	Q8K300	Q8K300 mus musculu
2	2313	57.3	1061	5	Q9V853	Q9V853 drosophila
3	1910	47.3	376	13	Q98784	Q98784 xenopus lae
4	1668.5	41.3	355	11	Q8BSC0	Q8BSC0 mus musculu
5	1552	38.4	288	4	Q96DE7	Q96DE7 homo sapien
6	1418	35.1	786	3	Q14366	Q14366 schizosacch
7	1398	34.6	911	4	Q8N5A7	Q8N5A7 homo sapien
8	1386	34.3	955	4	Q96PU5	Q96PU5 homo sapien
9	1384	34.3	947	4	Q7Z5F1	Q7Z5F1 homo sapien
10	1379	34.2	295	4	Q8NDP8	Q8NDP8 homo sapien
11	1378	34.1	258	11	Q9C853	Q9C853 mus musculu
12	1371	34.0	975	4	Q7Z5N3	Q7Z5N3 homo sapien
13	1369	33.9	967	4	Q7Z5F2	Q7Z5F2 homo sapien
14	1364.5	33.8	820	4	Q9NT88	Q9NT88 homo sapien
15	1357.5	33.6	855	11	Q8CF10	Q8CF10 mus musculu
16	1354.5	33.5	854	4	Q9H2W4	Q9H2W4 homo sapien

17	1354	33.5	835	11	Q8BRT9	Q8BRT9 mus musculu
18	1353.5	33.5	995	4	Q43165	Q43165 homo sapien
19	1353	33.5	855	4	Q8WU99	Q8WU99 homo sapien
20	1353	33.5	858	4	Q9BW58	Q9BW58 homo sapien
21	1346.5	33.3	855	11	Q99PK2	Q99PK2 mus musculu
22	1345.5	33.3	971	13	Q42573	Q42573 xenopus lae
23	1341.5	33.2	834	5	Q95RE4	Q95RE4 drosophila
24	1333	33.0	945	5	Q9YOH4	Q9YOH4 drosophila
25	1303.5	32.3	955	5	Q8IOR6	Q8IOR6 drosophila
26	1297	32.1	1007	5	Q9V7J3	Q9V7J3 drosophila
27	1274.5	31.6	792	5	Q95XU3	Q95XU3 caenorhabdi
28	1271.5	31.5	794	5	Q9N2T7	Q9N2T7 caenorhabdi
29	1244	30.8	728	4	Q9BKW4	Q9BKW4 caenorhabdi
30	1224.5	30.3	834	5	Q95TQ0	Q95TQ0 drosophila
31	1220.5	30.2	518	5	Q8TQC8	Q8TQC8 drosophila
32	1092	27.0	1581	4	Q9P2P5	Q9P2P5 homo sapien
33	1050	26.0	1585	4	Q9HCT7	Q9HCT7 homo sapien
34	1046.5	25.9	671	3	Q9UTG2	Q9UTG2 schizosacch
35	1032	25.6	222	11	Q80TB1	Q80TB1 mus musculu
36	1011.5	25.0	1583	11	Q8K4P8	Q8K4P8 mus musculu
37	926.5	22.9	437	5	Q9W3Z5	Q9W3Z5 drosophila
38	877.5	21.7	339	11	Q8B1A6	Q8B1A6 mus musculu
39	859	21.3	403	11	Q8VD14	Q8VD14 mus musculu
40	859	21.3	444	11	Q7TWI8	Q7TWI8 mus musculu
41	859	21.3	1080	11	Q921M5	Q921M5 mus musculu
42	859	21.3	1484	11	Q8CFH2	Q8CFH2 mus musculu
43	859	21.3	1906	4	Q15029	Q15029 homo sapien
44	859	21.3	3360	4	Q8NG67	Q8NG67 homo sapien
45	858	21.2	3227	3	Q13834	Q13834 schizosacch

ALIGNMENTS

RESULT 1
ID Q8K300 PRELIMINARY; PRT; 728 AA.
AC Q8K300;
DT 01-OCT-2002 (TREMBLrel. 22, Created)
DT 01-OCT-2002 (TREMBLrel. 22, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE Similar to E3 ubiquitin ligase SMURF1.
GN 4930431F1ORIK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_Taxid=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Strausberg R.;
RL Submitted (May-2002) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: CONTAINS 1 C2 DOMAIN.
DR EMBL; BC029097; AAH29097.1; -.
DR MGD; MGI:1923038; 4930431E1ORIK.
DR GO; GO:0005622; C:intracellular; IEA.
DR GO; GO:0016874; F:ligase activity; IEA.
DR GO; GO:0004842; F:ubiquitin-protein ligase activity; IEA.
DR GO; GO:0006512; P:ubiquitin cycle; IEA.
DR InterPro; IPR000008; C2.
DR InterPro; IPR008973; C2_CaB.
DR InterPro; IPR000569; HECT domain.
DR InterPro; IPR01202; WW_Rsp5_WWP.
DR Pfam; PF00168; C2; 1.
DR Pfam; PF00632; HECT; 1.
DR Pfam; PF00397; WW; 2.
DR SMART; SM00239; C2; 1.
DR SMART; SM00119; HECTC; 1.
DR SMART; SM00456; WW; 2.
DR PROSITE; PS00499; C2_DOMAIN_1; 1.
DR PROSITE; PS00004; C2_DOMAIN_2; 1.
DR PROSITE; PS02037; HECT; 1.
DR PROSITE; PS01159; WW_DOMAIN_1; 1.
DR PROSITE; PS00020; WW_DOMAIN_2; 2.

KW LIGASE. 728 AA; 83083 KW; C355291B9D8AD757 CRC64;
 SEQ SEQUENCE
 Query Match 75.3%; Score 3041; DB 11; Length 728;
 Best Local Similarity 75.0%; Pred. No. 1.7e-229;
 Matches 572; Conservative 63; Mismatches 78; Indels 50; Gaps 7;

QY 1 MNNPERRRNG-FVKRLRTYLCAKAK VKDFFLLPDPKAVVVDGSGGCHSTDTYKNTDP 59
 DB 1 MNNPERRRNG-SIKRLRLVLCAKAKAKDFRLLPDPFAKIVDGGGCHSTDTYKNTDP 60
 QY 60 KNNQHYDLYIGKSDSVTSVMNKKIKKKGAGFGFCRLSLNAINRLKDTGYORLDPCK 119
 DB 61 KNSQHYDLYIGKSDSVTSVMNKKIKKKGAGFGFCRLSLNAINRLKDTGYORLDPCK 120
 QY 120 LCPNNDYRGQIVVSLGSRDRIGTGGQVDCSRFLPNDLPDGMEEERTASGRIOYLNHI 179
 DB 121 LNPSTLDVARGQIVVSLGSRDRIGTGGQVDCSRFLPNDLPDGMEEERTASGRIOYLNHI 163
 QY 180 TATTGWERPTRASYSFSGFRPLSCFVDENPTISGTNATG-----QSSDPLAE 230
 DB 164 -----EDSGRGRPLSCIMEEPAPYTDGTAAGGNCRFVESPDSODQLLV 209
 QY 221 RRVRSQRHNVW----SRTLHTPPDLPEGEYRORTTQGGVYFLHTQVSTWDPVPR 286
 DB 210 QRLRPVEYRGPLQTPQNRPHGHOSPELBEGEYRORTTQGGVYFLHTQVSTWDPVPR 269
 QY 287 DLSNNGEIGPLPBGWEIRNTANCRRVYFVDHNNRTTOFTTPR-SANLHLVLRNQNLKD 346
 DB 270 DLSNNGEIGPLPBGWEIRNTANCRRVYFVDHNNRTTOFTTPR-SANLHLVLRNQNLKD 325
 QY 347 QQQQ-QVVSLCPDDTECTVPRYKRDLYQKLIKROELSDQOQAGCRISVSEETEE 405
 DB 326 PSQPLQPSSEGSVEDELPQORYERDLYQKLIKRLHLSLQDQAGCRISVSEETEE 385
 QY 406 SYRQMKRPDMDKRLMKIKRGEGLDYGVAEMTYLSHEMLNRYTG-FQYSRDIT 465
 DB 386 SYRQMKRPDMDKRLMKIKRGEGLDYGVAEMTYLSHEMLNRYTG-FQYSRDIT 445
 QY 466 TLQINPDSAVNPEHSEYFHFVGRIMGAVFPGHYIDGFTLPFYKOLGKSITLDDMELV 525
 DB 446 TLQINPDSINPDHLSYHFVGRIMGAVFPGHYIDGFTLPFYKOLGKSITLDDMELV 505
 QY 526 DPLHNSLVWLENDITGVLDHTFCVHNAYGEIIOHELKPKNSKIFVNEENKKEYRLY 585
 DB 506 DPLHNSLVWLENDITGVLDHTFCVHNAYGEIIOHELKPKNSKIFVNEENKKEYRLY 565
 QY 586 VNNRFLRGIZEQFALQGENEVIPOHLKTFDEKELIILICGLKIDVNDKVNTRLK 645
 DB 566 VNNRFLRGIZEQFALQGENEVIPOHLKTFDEKELIILICGLKIDVNDKVNTRLK 625
 QY 646 CTSPDSNIVKFWKAVEFFDEERRARLLQVYTGSSRVBLQGFKALQGAAGRLLFTIHQIDA 705
 DB 626 CTSPDSNIVKFWKAVEFFDEERRARLLQVYTGSSRVBLQGFKALQGAAGRLLFTIHQIDA 685
 QY 706 CTNNLPKRAHGTGNRIDIPIYESYKLYEKLLTAEECGRAVE 748
 DB 686 NTNNLPKRAHGTGNRIDIPIYESYKLYEKLLTAEECGRAVE 728

RESULT 2
 ID Q9V853 PRELIMINARY; PRT; 1061 AA.
 AC Q9V853 Q9U3W2;
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)
 DT 01-OCT-2002 (TrEMBLrel. 22, last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, last annotation update)
 DE CG94943 protein (E3 ubiquitin ligase) (Ubiquitin-protein ligase) (Smad-ubiquitin E3 ligase Smurf1).
 GN IACK OR SMURF OR SMURF1 OS CG94943.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;

OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_Taxid=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 PC STRAIN=Berkley.
 RX MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA Abail J.F., Agbayani A., An H.-U., Andrews-Framkoon C., Baldwin D.,
 RA Bailew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Bertman B.P., Bhattacharya D., Bolshakov S.,
 RA Burks D., Botchan M.A., Bouck J., Brockstein P., Broctier P.,
 RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.B., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
 RA Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glaeser K.,
 RA Glodok A., Gong P., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kamison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Iasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Izu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merklov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Paclow J.M.G.,
 RA Palazolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Klamas I., Simpson M., Skupski K.P., Smith T.,
 RA Spier E.C., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Switzkas R., Tector R., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,
 RA Williams S.M., Woodard T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of Drosophila melanogaster.";
 RL Science 287:2185-2195(2000).
 [2]
 RP SEQUENCE FROM N.A.
 RA Celniker S.E., Adams M.D., Kronmiller B., Wan K.H., Holt R.A.,
 RA Evans C.A., Gocayne J.D., Amanatides P.G., Brandon R.C., Rogers Y.,
 RA Banton J., An H., Baldwin D., Banton J., Beeson K.Y., Busam D.A.,
 RA Carlson J.W., Center A., Champe M., Davenport L.B., Dietz S.M.,
 RA Dodson K., Dorsett V., Doup L.E., Doyle C., Dresnek D., Farfan D.,
 RA Ferreira S., Frise E., Galie R.F., Garg N.S., George R.A.,
 RA Gonzalez M., Houck J., Hoskins R.A., Hostin D., Howland T.J.,
 RA Ibegwam C., Jalali M., Kruse D., Li P., Mattei B., Moshrefi A.,
 RA McIlroy T.C., Moy M., Murphy B., Nelson C., Nelson K.A., Nuno J.,
 RA Paclow J., Paragass V., Park S., Patel S., Pfeiffer B.,
 RA Phouanavong S., Pittman G.S., Puri V., Richards S., Scheeler F.,
 RA Stapleton M., Strong R., Switzkas R., Tector C., Tyler D.,
 RA Williams S.M., Zaveri J.S., Smith H.O., Venter J.C., Rubin G.M.;
 RT "Sequencing of Drosophila melanogaster genome.";
 RL Submitted (MAR-2000) to the EMBL/Genbank/DBJ databases.
 [3]
 RP SEQUENCE FROM N.A.
 RA Mitra S., Crosby M.A., Matthews B.B., Bayraktaroglu L., Campbell K.,
 RA Hradecky P., Huang Y., Kaminker J.S., Prochnik S.E., Smith C.D.,
 RA Tupy J.L., Bergman C., Bertman B., Carlson J.W., Celniker S.E.,
 RA Clamp M., Drysdale R., Emmert D., Frise E., de Grey A., Harris N.,
 RA Kronmiller B., Marshall B., Millburn G., Richter J., Russo S.,
 RA Seattle S.M.J., Smith E., Shu S., Smutnick F., Whitfield E.,
 RA Ashburner M., Gelbart W.M., Rubin G.M., Murgall C.J., Lewis S.E.;
 RA "Annotation of Drosophila melanogaster genome.";
 RT Submitted (MAR-2000) to the EMBL/Genbank/DBJ databases.
 [4]

RP SEQUENCE FROM N.A.
 RA Adams M.D., Celniker S.E., Gibbs R.A., Rubin G.M., Venter C.J.,
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 RN [5]
 RP SEQUENCE FROM N.A.
 RA Flybase;
 RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
 RN [6]
 RP SEQUENCE FROM N.A.
 RA Laurenceon A., Hawley S.,
 RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
 RN [7]
 RP SEQUENCE FROM N.A.
 RA Podos S.D., Hanson K.K., Wang Y.-C., Ferguson E.L.,
 RL "The Dmurt ubiquitin-protein ligase restricts BMP signaling spatially
 and temporally during Drosophila development.";
 Dev. Cell 1:0-0 (2001).
 RN [8]
 RP SEQUENCE FROM N.A.
 RA Liang Y.-Y., Lin X., Feng X.-H.,
 RL "dsurfi, a Smad-ubiquitin E3 ligase, specifically targets dpp-
 activated Mad protein for degradation.";
 Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
 CC -1- SIMILARITY: CONTAINS 1 C2 DOMAIN.
 DR EMBL; AEO03802; AAF57824.3; -;
 DR EMBL; AF216521; AAF21125.1; -;
 DR EMBL; AF416571; AAF09691.1; -;
 DR EMBL; AF464851; AAF09646.1; -;
 DR HSP; Q13526; IPIN.
 DR Flybase; FBgn0029006; lack.
 DR GO; GO:0005622; C:intracellular; IEA.
 DR GO; GO:0016874; E:ligase activity; IEA.
 DR GO; GO:0004842; F:ubiquitin-protein ligase activity; IEA.
 DR GO; GO:0006512; F:ubiquitin cycle; IEA.
 DR InterPro; IPR000973; C2_CalB.
 DR InterPro; IPR000569; HECT domain.
 DR InterPro; IPR001202; ww_Rsp5_wmp.
 DR Pfam; PF00168; C2; 1.
 DR Pfam; PF00632; HECT; 1.
 DR Pfam; PF00397; WW; 3.
 DR SMART; SMO0239; C2; 1.
 DR SMART; SMO0456; WW; 3.
 DR PROSITE; PSS0049; C2_DOMAIN_1; 1.
 DR PROSITE; PSS0004; C2_DOMAIN_2; 1.
 DR PROSITE; PSS0237; HECT; 1.
 DR PROSITE; PSS0159; ww_DOMAIN_1; 2.
 DR PROSITE; PSS0020; ww_DOMAIN_2; 3.
 DR Ligase.
 KM
 SQ SEQUENCE 1061 AA; 115675 MW; 68BC550F5129163 CRC64;

Query Match 57.3%; Score 2313; DB 5; Length 1061;
 Best Local Similarity 45.1%; Pred. No. 3.5e-172;
 Matches 476; Conservative 107; Mismatches 157; Indels 316; Gaps 15;

QY 7 RANGPVKRLTLVLCARLVKDFRLPDPFAKVVVDGSGCHSTDTVKNLTIDPKMNOHYD 66
 DB 8 RRGTHVVRITILCARLVKDFRLPDPFAKVVVDGSGCHSTDTVKNLTIDPKMNOHYD 67

QY 67 LYIKSGSVTSVWNHKKIKKQAGFLGCVRLISNAINLKQNGVRLDLCTGSPDND 126
 DB 68 LFLDIGAIIITTVNOKRIHK-GSGFLGCVRFPAFNIQSLKAGFQRLDLGKLSPPDD 125

QY 127 TVSGQIVSLQSRD-----IGTGQGVVDCSR--LFDNDLPDGMERRRTASGRIOY 175
 DB 126 LVKQIITISLSSKQPSGNGNLAVGSGDVKPSDDSSDLSLPEMWEERTNGKVVY 185

QY 176 LNHITRTTQWERFPR----- 191
 DB 186 VNHATKSTQWDRPQPGVGVSSSHATSPOQRHNTNGNSGDRQADAGPTSTTTCTNLMNG 245

QY 192 -----ASEVSPGRPLSCFV----- 206

DB 246 HRSRDLSTVTSADRRRSTELLSSVGKENTSPPTPVGATTTGPKKTSSNSAGRTLEQ 305
 QY 207 -----DENTPIGCT-----N 216

DB 306 RPTNEPATPTSTTSASVRLHNDNHVKTPKQTNHAPPESTPTSGQGVNNGNAQN 365
 QY 217 GAT-----CGQSSDPLAER----- 231

DB 366 GSTSGSGGQAAPQASNSWTOEDAAITTSPTSTTSPPRHSGSPPTPNISPPASVPSA 425
 QY 232 -----RVRSQRHRYNSRTHLTPP----- 251

DB 426 NGNVHSPNANSTPAGSGGSRSYTATPGQRQRSSRQGESESTERRSSRGTNGTS 485

QY 252 -----DIPGEQRTTGGGVYFLHTQVSTWHDPR 283

DB 486 GGGGGGGGQRYASAAIAAQAAPPLDPPGTEMKRTQQGGVYFHTQVSTWHDPR 545

QY 284 VPRDL--SNINCEELGPLPGEWEIRNTATGRVYFVDHNNRTTOFTDRLSANLVLNRO 341
 DB 546 IPRPDYQHLTLDAIGLPSGWEQRKTAAGRVYVDHNNRTTOFTDRLSGSILQMRG 605

QY 342 N-----QLKQOQOQVSL-----CPDTE 361

DB 606 TVPPTSANAGTPAPPSATPAPSAAPPAATPASNATPTTLTTNPPHRIVPLPQ 665

QY 362 CL-----TVPRYKRDLYOKLTLRQESLQOOPQAGHCRIEVSREIFEESYQVMKRPX 416

DB 666 GLRGADLIPKRYRDLVNGKRALMTLEIQWQPSGCHRLTQVSRNREIFEESYRLMKRAK 725

QY 417 DLMKRLMKRGEGLDYGVAKEMLYLHSEMLNPPYGLFOYSRDDITTYQINPDAVN 476

DB 726 DMKRLMKVFKGEGLDYGVAKEMLYLHSEMLNPPYGLFOYSRDDITTYQINPDAVN 785

QY 477 PEHLSTYFHFVGRINGMAVFGHYIDGGFTLPFYQLLGKSTTLDMELVBDLHNSLVI 536

DB 786 PDHLSYFHFVGRINGMAVFGHYIDGGFTLPFYQLLGKSTTLDMELVBDLHNSLVI 845

QY 537 LENDITGVLDHTFCVENHAYEIIQHELPNGKSIPIVNEENKEVYVLYNMRFLRGIEA 596

DB 846 LENSISGLIESTFVSENNISFGALVYVHLEKPGASIPVTEENKREYKLYVYRFRGIEQ 905

QY 597 QFLALQKRPNYIPOHLIKTDEKLELIIGLGLKIDVDMKNAITKCHCPDSNIVKWF 656

DB 906 QFLALQKFCFELIPHLRPFDERLELVIGISISIDVDMKNAITKCHCPDSNIVKWF 965

QY 657 WKAVEFPDEERRARLQFVTGSSRVPLQGFALQ---GAAGPRLFTTH-QIDACTNNLPK 712

DB 966 WQVYESISEMRAKRLQVGTSSRVPLQGFALQSGTAVGPRFLTTHLTADVPLQNLPK 1025

QY 713 AATCFNRIDIPPESEYKLYEKLTAIEETGFAVE 748

DB 1026 AATCFNRIDIPPEYTYQLCDKLTQAVEETGFAVE 1061

RESULT 3
 Q98TS4 PRELIMINARY; PRT; 376 AA.
 AC Q98TS4;
 DT 01-JUN-2001 (TrEMBLrel. 17, Created)
 DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE E3 ubiquitin ligase Smurf2 (Fragment).
 OS Xenopus laevis (African clawed frog).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
 OC Xenopodinae; Xenopus.
 OC NCBI_Taxid=8355;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21107656; PubMed=1158580;
 RA Zhang Y., Chang C., Gehring D.J., Hemmati-Brivanlou A., Derynck R.;

RT "Regulation of Smad degradation and activity by Smurf2, an E3
 RT ubiquitin ligase.";
 RL Proc. Natl. Acad. Sci. U.S.A. 98:974-979(2001).
 CC -1- SIMILARITY: CONTAINS 1 C2 DOMAIN.
 DR EMBL; AY014181; AAG50422.1; -.
 DR HSSP; O13526; 1PIN.
 DR GO; GO:0016874; F:ligase activity; IEA.
 DR InterPro; IPR000008; C2.
 DR InterPro; IPR008973; C2_GaIB.
 DR InterPro; IPR001202; WW_Rsp5_WWP.
 DR Pfam; PF00168; C2; 1.
 DR Pfam; PF00397; WW; 3.
 DR SMART; SM00239; C2; 1.
 DR SMART; SM00456; WW; 3.
 DR PROSITE; PS00004; C2 DOMAIN 2; 1.
 DR PROSITE; PS01159; WW DOMAIN 1; 1.
 DR PROSITE; PS0020; WW_DOMAIN 2; 3.
 DR Ligase.
 KW NON_TER
 FT SEQUENCE 376 AA; 42652 MW; CA6AB2B5F4C20F98 CRC64;
 Query Match 47.3%; Score 1910; DB 13; Length 376;
 Best Local Similarity 94.1%; Pred. No. 3e-141;
 Matches 354; Conservative 10; Mismatches 8; Indels 4; Gaps 3;
 QY 1 MSNPRRRNGPVKRLTVLCAKNIWKDFR-LPPFAKVVVDGSGGCHSTDTYKNTLP 59
 DB 1 MSNOSRRNGPVKRLTVLCAKNIWKDFRGLTLDSPAKVVVDGSGGCHSTDTYKNTLP 60
 QY 60 KMNQHYDLYIGKSSVTSVWNNKKIKHKQAGLGCVRLLSNAINFLKDTGYQRDLCK 119
 DB 61 KMNQHYDLYIGKSSVTSVWNNKKIKHKQAGLGCVRLLSNAINFLKDTGYQRDLCK 120
 QY 120 LGPNDNTPVRCQIVLSLQSRDRICTGGGVVDCSLFNDLPDGEERTASGRIQVLYNH 179
 DB 121 LGPNDNTPVRCQIVLSLQSRDRICTGGGVVDCSLFNDLPDGEERTASGRIQVLYNH 180
 QY 180 TRTTQMBEPTRPASEYSSGPRPLSCFYDENTPIISGTNGATCGOSSDRLAEREVRQRRH 239
 DB 181 TRTTQMBEPTRPASEYSSGPRPLSCFYDENTPIISGTNGATCGOSSDRLAEREVRQRRH 240
 QY 240 NYMSRTHLHPEDLPBEGYEORTTQGGVVFLLHTQVSTVHDPRVPRDLSNINCEBGL 299
 DB 241 NYMSRTHLHPEDLPBEGYEORTTQGGVVFLLHTQVSTVHDPRVPRDLSNINCEBGL 300
 QY 300 PPGNEIRNTATGRVYFVDHNNRTQFTDPRLSANLHLVLRN-OLKDOQQQOVSLC-- 356
 DB 301 PPGNEIRNTATGRVYFVDHNNRTQFTDPRLSANLHLVLRN-OLKDOQQQOVSLCQL 360
 QY 357 PPDTECLTVRRYRDL 372
 DB 361 PDEVECLTVRRYRDL 376
 RESULT 4
 Q8BSC0 ID Q8BSC0 PRELIMINARY; PRT; 355 AA.
 AC Q8BSC0;
 DT 01-MAR-2003 (TREMBlrel. 23, Created)
 DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)
 DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
 DE Hypothetical HECT domain (Fragment).
 GN 4930431E10R1K.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxId=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Embryo;
 RX MEDLINE=2354683; PubMed=1246651;
 RA The PANTOM Consortium.
 the RIKEN Genome Exploration Research Group Phase I & II Team;

RT "Analysis of the mouse transcriptome based on functional annotation of
 RT 60,770 full-length cDNAs.";
 RL Nature 420:563-573(2002).
 DR EMBL; AK034736; BAC28813.1; -.
 DR MGD; MGI:1923038; 4930431E10R1K.
 DR GO; GO:0005622; C:intracellular; IEA.
 DR GO; GO:0004842; F:ubiquitin-protein ligase activity; IEA.
 DR GO; GO:0006512; P:ubiquitin cycle; IEA.
 DR InterPro; IPR000569; HECT_domain.
 DR Pfam; PF00632; HECT; 1.
 DR SMART; SM00119; HECTC; 1.
 DR PROSITE; PS0237; HECT; 1.
 KW Hypothetical protein.
 FT NON_TER
 FT SEQUENCE 355 AA; 41315 MW; 93B39B4C82F86DD1 CRC64;
 Query Match 41.3%; Score 1668.5; DB 11; Length 355;
 Best Local Similarity 86.5%; Pred. No. 2.4e-122;
 Matches 307; Conservative 25; Mismatches 20; Indels 3; Gaps 1;
 QY 397 VRESEIFESRYQVKKRPKDIWKGLMTKFGEBGLDYGVARWLYLSHEMLNPPYGL 456
 DB 1 VRESEIFESRYQVKKRPKDIWKGLMTKFGEBGLDYGVARWLYLSHEMLNPPYGL 60
 QY 457 FQYSRDDTYTQINPDSAVNPEHLSYFPRVGRIMGAVFHGHYIDGFTLFFYQLGKS 516
 DB 61 FQYSTDNITTYTQINPDSINPDHLSYFPRVGRIMGAVFHGHYINGGTFVYQLGSKP 120
 QY 517 ITLDMELVDPDLHNSLVLIENDITGVLDHTFCVENHAYEIIQHEIKPKNGKSI PVNEE 576
 DB 121 IQLSDLESVDPELHKSILVLIENDITPVLDHTFCVENHAFRIILOHEIKPKNGRAVPTEE 180
 QY 577 NKEEYVRLVYMKRPFRIEAGFLOKGFENVIQHLKTPDELELITGLKIDVND 636
 DB 181 NKEEYVRLVYMKRPFRIEAGFLOKGFENVIQHLKTPDELELITGLKIDVND 240
 QY 637 WKQNTRLXGCTPDSNIVKFWKAVEFPDEERRARLLQVTSSSVPLQGFALQ---GAA 693
 DB 241 WKSNTRLXGCTPDSNIVKFWKAVEFPDEERRARLLQVTSSTVPLQGFALQGSFGAA 300
 QY 694 GPRLFTTHQIDACTNLPKATCTNRIDIPYSEYKLYEQLTAIBETGFAVE 748
 DB 301 GPRLFTTHQIDACTNLPKATCTNRIDIPYSEYKLYEQLTAIBETGFAVE 355
 RESULT 5
 Q96DE7 ID Q96DE7 PRELIMINARY; PRT; 288 AA.
 AC Q96DE7;
 DT 01-DEC-2001 (TREMBlrel. 19, Created)
 DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
 DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
 DE Similar to E3 ubiquitin ligase SMURF2 (Fragment).
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxId=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC Tissue=Lung;
 RA Straubeberg R.;
 RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; BC009527; AA09527.1; -.
 DR GO; GO:0005622; C:intracellular; IEA.
 DR GO; GO:0016874; F:ligase activity; IEA.
 DR GO; GO:0004842; F:ubiquitin-protein ligase activity; IEA.
 DR GO; GO:0006512; P:ubiquitin cycle; IEA.
 DR InterPro; IPR000569; HECT_domain.
 DR Pfam; PF00632; HECT; 1.
 DR SMART; SM00119; HECTC; 1.
 DR PROSITE; PS0237; HECT; 1.
 KW Ligase.
 FT NON_TER
 FT SEQUENCE 1 1

SQ SEQUENCE 288 AA; 33255 MW; FE2B43E300E66537 CRC64;
Query Match 38.4%; Score 1552; DB 4; Length 288;
Best Local Similarity 100.0%; Pred. No. 2,4e-113; Indels 0; Gaps 0;
Matches 288; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 461 RDDIYTLQINPDSAVNPEHLSYFHEVGRIMGAVFHGHYIDGGFTLPFYKQLIGSITLD 520
DB 1 RDDIYTLQINPDSAVNPEHLSYFHEVGRIMGAVFHGHYIDGGFTLPFYKQLIGSITLD 60
QY 521 DMEIYVDPDLNSIYVLTLENDITGVLDHTECVENHNAAGELIOHEIKPNKSGSIYVNEENKE 580
DB 61 DMEIYVDPDLNSIYVLTLENDITGVLDHTECVENHNAAGELIOHEIKPNKSGSIYVNEENKE 120
QY 581 YVRLYVNWRFELRGIEAOFALQKGFNEVIAPOHLKTFDEKELELIIGLAKIDVNDKVN 640
DB 121 YVRLYVNWRFELRGIEAOFALQKGFNEVIAPOHLKTFDEKELELIIGLAKIDVNDKVN 180
QY 641 TRLKCTPDGSIYKMFKAWEFDEERRARLLQFVTGSSRVPLQGFKAQAGAPRLFTI 700
DB 181 TRLKCTPDGSIYKMFKAWEFDEERRARLLQFVTGSSRVPLQGFKAQAGAPRLFTI 240
QY 701 HQIDACTNNLPKATCFNRIDIPYSEYEKLYEKLTAIEBTGFAVE 748
DB 241 HQIDACTNNLPKATCFNRIDIPYSEYEKLYEKLTAIEBTGFAVE 288

RESULT 6
014326 PRELIMINARY; PRT; 786 AA.

ID 014326 AC
DT 01-JAN-1999 (TEMBLrel. 09, Created)
DT 01-JUN-1999 (TEMBLrel. 09, Last sequence update)
DT 01-OCT-2003 (TEMBLrel. 25, Last annotation update)
DE Ubiquitin--Protein Ligase.
GN SPBC169.11C.
OS Schizosaccharomyces pombe (Fission Yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
OC NCBI_TaxID=4896;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=972h-;
RA Volckert G., Wood V., Rajandream M.A., Bartell B.G.;
RL Submitted (Aug-1997) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: CONTAINS 1-C2 DOMAIN.
DR EMBL; Z99759; CAB16903.1; -;
DR PIR; T39585; T39585.
DR HSP; Q13526; PIN.
DR GenDB Sprobe; SPBC169.11C; -;
DR GO; GO:0005682; C:intracellular; IEA.
DR GO; GO:0016874; F:ligase activity; IEA.
DR GO; GO:0004842; F:ubiquitin-protein ligase activity; IEA.
DR GO; GO:0006512; P:ubiquitin cycle; IEA.
DR InterPro; IPR000008; C2.
DR InterPro; IPR008973; C2_CALB.
DR InterPro; IPR000569; HECT_domain.
DR InterPro; IPR003449; WW.
DR InterPro; IPR001202; WW_reps_WWP.
DR Pfam; PF0168; C2; 1.
DR Pfam; PF00632; HECT; 1.
DR Pfam; PF00397; WW; 3.
DR PRINTS; P00403; WMDOMAIN.
DR SMART; SM00239; C2; 1.
DR SMART; SM00119; HECT; 1.
DR SMART; SW00456; WW; 3.
DR PROSITE; PS00499; C2_DOMAIN_1; 1.
DR PROSITE; PS00004; C2_DOMAIN_2; 1.
DR PROSITE; PS0237; HECT; 1.
DR PROSITE; PS0159; WW_DOMAIN_1; 3.
DR PROSITE; PS0020; WW_DOMAIN_2; 3.
KM Ligase.

SQ SEQUENCE 786 AA; 89259 MW; 57B7A859F5497B9A CRC64;
Query Match 35.1%; Score 1418; DB 3; Length 786;
Best Local Similarity 36.7%; Pred. No. 3.2e-102;
Matches 316; Conservative 127; Mismatches 217; Indels 202; Gaps 20;

QY 10 GPVTLRLTVCAKLVKKDFRLLDPFAKVVYVGGSGCHSDYKRLDPRKNDYLDYI 69
DB 4 GAKVRFYVAADLSRDLFRQDPFALLVDSG-OTHTTKYIKSVNPTNMGFEVY 62
QY 70 GKSDSVISVNNHKKIKHKGAGFLGCVRLSNAINELKDTGYORDL-CKLGP 122
DB 63 KPSSVISIRLFDQKFF-KKQDQFLGVSF--RMREVSFRSNREVSILRPLKSS 115
QY 123 NDNPTVAGQIVV----- 134
DB 116 TTNLSVIGNLVILKVAPSKIRAPAGNHSSTANRTSTPTTTATRTTRPATNTSN 175
QY 135 -----SLGSRDRIGTG------QVDCSRFLDND----- 158
DB 176 QSTNSTNGTSATSNCTGTGAGTGAHSRSPVTRQNTGSLSNMHNKSSFDQY 235
QY 159 --LPDGEERTASGRIOYLNHTTTQWERPTRPASEYSGRPLSCFVDENTPIGCTN 216
DB 236 GRLPGEWRRADSLGRTYVDHNTRTTW--TRPAS-----STNPVNT- 277
QY 217 GATGQSSDPRFLARRVRSORRRVMSRTHLTPP-----DLPEGYEORTQ 263
DB 278 -----SSD-----SQR-LVHQR--HLPDSSNPSSLNQSDGNDLPFGEMRYTD 318
QY 264 QGOVYFHTQTVGSIWMDPR-----VPRDLNINCEELGRLPGWEIR 306
DB 319 TGRPFYVDHNTRTTWVDFRNPVLRPNCGSSVTGSLMQPQLSH-----LGPLSGWEMR 373
QY 307 NTATGRVYFVDHNNNTTFTPRLSANHLVLRNQNLKQDQOQVSLCPDDTBECLTP 366
DB 374 LTNARVYFVDHNTTTWDDPRLPSAL-----DD-VP 406
QY 367 RYKRDLYVCKLILROELSQOQPOAGHCRIEVSRHEIFPESYQVKKRPKDLMKRLMKF 426
DB 407 QYKDFPRKLIYFRQ-PGMRPLPGQCNVYKRBDIIFEDSYAEIYRVAHDLKKRLMIRF 465
QY 427 RGEGLDYGVAREKLVYLSEHMLNPYGLFOYSEDDYITQINPDSAVNPEHLSYFHEV 486
DB 466 DEEDLDYVGLSRSEFFFLSHMPDPICLPEYSADVYTLQINPHSSINSEHLYFRFI 525
QY 487 GRIMGAVFHGHYIDGGFTLPFYKQLIGSITLDMEIYVDPDLHNSLWILENDITGILD 546
DB 526 GRVIGLAIFFRRFLDAFVVSILYKLLKRYSLADMSIDAEFYSLKVLENDITGILD 585
QY 547 HTFCVENHNAAGELIOHEIKPNKSGSIYVNEENKEVRLYVNWRFELRGIEAOFALQKFN 606
DB 586 LTFVSVEEDHFGVRVEVLTNGENIEVTEENKKYVDLVTEWRVSKRVEQGFNAYSGFV 645
QY 607 EVTPQHLKTPREKLELIIGLAKIDVNDKVNRLHCTPDSIYKMFKAWEFDEE 666
DB 646 ELVSPVLVNFDERLELIIGISVDVDMKSHREYTYIATDVAIWMFETLWGKXE 705
QY 667 RARLLQFVTGSSRVPLQGFKAQAGAPRLFTIHOIDACTNNLPKATCFNRIDIPYE 726
DB 706 DRSKLLQFATGSRIPVNGFRDLQSGDGRKRTIKA-GTPQLVATCFFRLDLPYV 764
QY 727 SYEKLYEKLTAIEBTGFAVE 748
DB 765 SKDTLHEKLSLAVENTVPGFNE 786

RESULT 7
08NSA7 PRELIMINARY; PRT; 911 AA.
ID 08NSA7 AC
DT 01-OCT-2002 (TEMBLrel. 22, Created)
DT 01-OCT-2002 (TEMBLrel. 22, Last sequence update)

01-OCT-2003 (Tremblrel. 25, Last annotation update)
 DE Similar to neural cell expressed, developmentally down-regulated
 DE 4-like
 OC Homo sapiens (Human)
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OC NCBI_Taxid=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Melanoma;
 RA Strusberg R.;
 RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
 CC -1- SIMILARITY: CONTAINS 1 C2 DOMAIN.
 DR EMBL; BC032597; AAH32597.1; -
 DR GO; GO:0005622; C:intracellular; IEA.
 DR GO; GO:0004842; F:ubiquitin-protein ligase activity; IEA.
 DR GO; GO:0006512; P:ubiquitin cycle; IEA.
 DR InterPro; IPR000008; C2_CaLB.
 DR InterPro; IPR008973; C2_CaLB.
 DR InterPro; IPR002349; WW.
 DR InterPro; IPR001202; WW_Rsp5_WWP.
 DR Pfam; PF00168; C2_1.
 DR Pfam; PF00632; HECT_1.
 DR Pfam; PF00397; WW_3.
 DR PRINTS; PR00360; C2DOMAIN.
 DR PRINTS; PR00403; WMDOMAIN.
 DR SMART; SM00239; C2_1.
 DR SMART; SM00119; HECTC_1.
 DR SMART; SM00456; WW_3.
 DR PROSITE; PS00049; C2_DOMAIN_1; 1.
 DR PROSITE; PS50237; HECT_1.
 DR PROSITE; PS01159; WW_DOMAIN_1; 3.
 DR PROSITE; PS50020; WW_DOMAIN_2; 3.
 SQ SEQUENCE 911 AA; I04921 MW; CE04AAED677AA506 CRC64;

Query Match 34.6%; Score 1398; DB 4; Length 911;
 Match Local Similarity 35.18; Pred. No. 1.5e-100;
 Matches 327; Conservative 123; Mismatches 236; Indels 246; Gaps 24;

14 LRLTLVAKNLYKDFRLPDPFAKV--VVDGSGGCH--STDVTKTLDPKKNQNDY 68
 22 LRVAVGIDAKDKITIGASDPYKLSLVADENRELALVOTIKITLNPKNNEFYR 81
 69 ICKSD-SVTTISVNNHKKIKKKQAGPLGVR-----LTSNAIN 105
 82 VNPENHRLPEVFPDENLRTD--FLGQVDPVLSHLPTEPTMERPYTEKDLRPRSH 138
 106 RLKDTGYQSLDLCKLGFNDNDVYRGQIVSLQSDR:GTGGQVVDGSRLEND----- 158
 139 KSRVKGFLRLKMAVMPNG-----GDDENSDQSDVHEGWEVVD-----SNDASQHOE 188
 159 -----LPDGEERTASGRIOYLNHITRTTOWERPT----- 189
 189 ELPPPLPPGWEKEVDMLGRYYVNNHRTTQWHPFLMDVSESDNNIRQINQZAHBR 248
 190 -----RPAPE-----YSSFG----- 199
 249 FRSRHHSIDLEPPSGGVPEPWETISEVINIAGDSLGLALPPPAASGTSFQELS 308
 200 -----RPLSGFVDENT-----RPLSGTNGATGCGS 224
 309 EELSRRLQITPDNSNGEFGSLIQBPSSRLRSCSVTDVAVEQHLPLMADGASGATNS 368
 225 DPLRAERVASQRRHNMST----- 245
 369 NNHILFQIR--RPRSSPTVTLASLEGAKDSPVRAVKDTLSNPQSPFPYNSPKR 426
 246 -HLATPDLDEGEYEQRTTQGOVYFLHTGTGVSTWHDPRVP-----RDLSTNICEELGFL 299
 427 QHATYGSFLPPGWEMLRPAWGRPFIFDHNKITTWEDPRKXFVHNRKSTSLNPNDLGFL 486

300 PGWEIRNTATGRVYFVNNHNTTQPTDPRLSANHLVLTNQNOLKDDQOOQVSLCPDD 359
 487 PGWEIRHILDRITIDHNSKITQWEDPL-----QN----- 519
 360 TECTLTPR--YRDLVQKLIKIRQELSQOQAGCRILEVREELFEESYQVMKRPKD 417
 520 -PAITGPAPVYRERKQKDYPRKXKLPADIPNFEKKHRRNNIFESYRIMS VKRPD 578
 418 LMK-RLMTKFRGEEGLDGVAREWLYLSHMLNPYYGLQYSKDDIYTIQINPDSAV- 475
 579 VKARLMTLFESEKGLDGVAREWFLPSKMFNPYYGLFESATDWTYTIQINPDSLC 638
 476 NPEHLSYFHFVGRIMGAVFEGHYIDGFTLPFYQLGKSIITLDDMELVDPDLNLSLV 535
 639 NEDHLSYFFIFIRVAGLVAFHCKLIDGFIRPFYQMLCKQITLNDMSVSEYNSLKW 698
 536 ILENDITGVLDHFCVEHNAVEIIOHLPKNGKSPVNEENKKEYVLVNNMRPLRGIE 595
 699 ILENDPT-ELDLMFCIDENFEQYQVLDKNGSEIWTNENKKEYIDLVIQWFRVNVQ 757
 596 AQLFLOGFNEVIFQHLKTFDEKELILICGLKIDVNMKXNTRLK--HCTPDSNTV 653
 758 KQMAAFLEGFTLPLIDILIKI FDENELBLMCGLDVDVNMQRSHYKNGYC-PNHEVI 816
 817 QMFKAVALIMDAEKIRLLQFYGTGRVPMNGFAELVGSNGPQFTIEQWGS-DEKLPRA 875
 714 HTCFNRIDIPYSEYKYEKLTIAETCP 745
 876 HTCFNRIDIPYETEDREKLMVAENAGQF 907

RESULT 8
 096FUS PRELIMINARY; PRT; 955 AA.
 ID 096FUS;
 AC 01-DEC-2001 (Tremblrel. 19, Created)
 DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)
 DT 01-OCT-2003 (Tremblrel. 25, Last annotation update)
 DE NEDD4-like ubiquitin ligase 3.
 GN NEDL3.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OC NCBI_Taxid=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Okamoto Y., Miyazaki K., Sakamoto M., Kato C., Nakagawara A.;
 RT "Homo sapiens NEDD4-like ubiquitin ligase 3";
 RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RA MEDLINE=21269431; PubMed=11244092;
 RT Harvey F.K., Dinudom A., Cook I.D., Kumar S.;
 RL "The Nedd4-like protein KIAA0439 is a potential regulator of the epithelial sodium channel";
 RL J. Biol. Chem. 276:8597-8601 (2001).
 RN [3]
 RP SEQUENCE FROM N.A.
 RA MEDLINE=20501262; PubMed=11046148;
 RX Winberg G., Matskova L., Chen F., Plant P., Rotin D., Gish G.,
 RA Ingberg R., Ernberg I., Pawson T.;
 RL "Latent membrane protein 2A of Epstein-Barr virus binds WW domain E3 protein-ubiquitin ligases that ubiquitinate B-cell tyrosine kinases";
 RL Mol. Cell. Biol. 20:8526-8535 (2000).
 CC -1- SIMILARITY: CONTAINS 1 C2 DOMAIN.
 DR EMBL; AB071179; BAB69424.1; -
 DR GO; GO:0005622; C:intracellular; IEA.
 DR GO; GO:0016874; F:ligase activity; IEA.
 DR GO; GO:0004842; F:ubiquitin-protein ligase activity; IEA.
 DR GO; GO:0006512; P:ubiquitin cycle; IEA.
 DR InterPro; IPR000008; C2.

```

DR InterPro; IPR008973; C2_CaB.
DR InterPro; IPR000569; HECT_domain.
DR InterPro; IPR002349; WW.
DR InterPro; IPR001202; WW_Rsp5_WWP.
DR Pfam; PF00168; C2; 1.
DR Pfam; PF00632; HECT; 1.
DR Pfam; PF00397; WW; 4.
DR PRINTS; PR00360; C2DOMAIN.
DR PRINTS; PR00403; WWDOMAIN.
DR SMART; SM00239; C2; 1.
DR SMART; SM00119; HECTC; 1.
DR SMART; SM00456; WW; 4.
DR PROSITE; PS00499; C2_DOMAIN_1; 1.
DR PROSITE; PS00004; C2_DOMAIN_2; 1.
DR PROSITE; PS02337; HECT; 1.
DR PROSITE; PS01159; WW_DOMAIN_1; 4.
DR PROSITE; PS00202; WW_DOMAIN_2; 4.
DR LIGASE.
KW LIGASE.
SQ SEQUENCE 955 AA; 110021 MW; A8BB278A37F6A6B5 CRC64;

Query Match 34.3%; Score 1386; DB 4; Length 955;
Best Local Similarity 34.4%; Pred. No. 1,4e-99;
Matches 330; Conservative 125; Mismatches 247; Indels 258; Gaps 26;

14 LRLTVLCAKNLVKDFRLLPDPFAKV---VVDGSGGCH--STDVYKNTLDPKNGHYDLY 68
22 LRVKVVSGIDLAKDIFGASDPYKLSLYVDENRRLALVQTKTKTLNPKNNEFFYR 81
69 IGKSD-SVTISVNNHKKHKQAGFLGCVR-----LISNAIN 105
82 VNPNSHRLLFVYDENRRLTRDD---FLGOVDP.LSHLPTEDPTMERPYTFKDFLLRPSH 138
106 RLKDTGYORLDCKLGPNDNDTVRGQIVVLSQSRDILGTGGVDCSRLLFNDN----- 158
139 KSRVKGFLRLKMAVFKNG-----GODENSDQRDMDHEGWEEVD-----SNDASQHOE 188
159 -----LPDGMERRRTASGRIOYLNHITRTTOWERPT----- 189
189 ELPPPLPFGMEBKVDNLGRITVYVNNHNRRTQWHRP.SLMDVSSDNNIRQINQEAHRR 248
190 -----RPASEYSSP-----GRPLSCFVDENPTISGNT-----GATCGQSSDPR-- 227
249 FRSRHHISDLEBPSEGGDVPEPMEWT-ISEEVNIAGDSLGLALPPPPASGSRTPSQEL 307
228 -----LAER-----RVRS-----QRHNNYSRTHLHTPPDL 253
308 SEELSRRLQITTPDSNGEQFSSLIQREPSRLRSCSVTDAVAEQGHLPPPSVAVHTTPEL 367
254 PEGYEQRTTOOGQVYFLHTQGVSTWHD-----RVPRDL 288
368 PSGWEEBKDAKGRITVYVNNHNRRTTWTPTMQLAEDGASGATSNNNHLIEPQIRRPRL 427
289 SNINCEELGP-----LPFGWEIN 307
428 SSGTVTISAPLEGAKDSVPRRAVKDTLSNPQSPQSPYNSPKPOHKYTGSLPPGWEKRI 487
308 TATGRVYFVDHNNRTTQFTDPRLSANLHLVLRNQLKDQOQOQVSLCPDD----- 359
488 APNGRPFFIDHNTKTWEDPRLKFPVHM-----RSKSLNPNNDGRLPPGM 534
360 -----TECLTVPR--YKRDVQLKTLRQSLQOQOQ 389
535 EERIHLDGRFFYVDHNSKITQWEDPRLONFATGPAVPSREPKQKYDYFRKDLKPKPAI 594
390 AGCRLEVSREELFEESYQVMKRPDLWK-RLMTKFRGEEGLDYGVAREWLYLSHE 448
595 PNFPEYKLRHNNIFEEZYRIMSVPKPDV.KARLMTIEFESEKGLDGVAREWFLFSKE 654
449 MNPYYLGLPOYSRNDITLQINPDSAV-NPEHLSTPHFVGRINGMVFCHYIDGGFTLP 507
655 MNPYYLGLPESATDNTLQINPNSGLCNEDHLSYTFIGRVAGLVFHEKLDGFFIR 714
508 FYKOLLGKSLTLDLMELVDPDLHNSLVLWLENDITGVLDHTFCVENNAVEGELLIOHEIKDN 567

```

```

DB 715 FYKMLGKQITLDMESVDSEYINSKMLLENDPT-ELDMFCIDENPFQATQVLDKPN 773
QY 568 GKSIPVNEENKKEYVRLVYNNRFLRGIEAQLQKGFNEVLPQHLKTFDEKELELIIC 627
DB 774 GSEIMVTNNKREYIDLVIQRFVNRVQKQNNAFLEGFTELLPDLIKIFDENLELDMC 833
QY 628 GLGKIDVNDKMNTRLK--HCTPDSNVKWEFKAVPEFDEERRARLLQFTGSSRPVLOG 685
DB 834 GLGSDVDVDMQSHSIYXGVC-PNHPIVQFMKAVLMLDMDEKRIRLQFTVGTSRVPNG 892
QY 686 FKALQAGAPRLFTIHQIDACTNNLPKAAHCFNKRIDIPYSEYEKLYEKLLTALETCGF 745
DB 893 FAELYGNGPOLFTIEQWGS-PEKLPRAHICFNRIDLPYETEPEDREKILMAVENNAQGF 951

RESULT 9
Q725F1 PRELIMINARY; FR; 947 AA.
AC Q725F1
DT 01-OCT-2003 (TrEMBLrel. 25. Created)
DT 01-OCT-2003 (TrEMBLrel. 25. Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25. Last annotation update)
DE Ubiquitin ligase NEDD4L.
GN NEDD4L.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OC NCBI_TaxId=9606;
RN (1)
RP SEQUENCE FROM N.A.
RA Qi H., Labrie C.;
RT "New splicing isoform of human Nedd4-2."
RL Submitted (MAY-2002) to the EMBL/Genbank/DBJ databases.
DR EMBL; AY12985; AAM76730.1; -.
KW LIGASE.
SQ SEQUENCE 947 AA; 109429 MW; 95CB0FDEDCB96639 CRC64;

Query Match 34.3%; Score 1384; DB 4; Length 947;
Best Local Similarity 34.4%; Pred. No. 2e-99;
Matches 330; Conservative 125; Mismatches 247; Indels 258; Gaps 26;

14 LRLTVLCAKNLVKDFRLLPDPFAKV---VVDGSGGCH--STDVYKNTLDPKNGHYDLY 68
14 LRVKVVSGIDLAKDIFGASDPYKLSLYVDENRRLALVQTKTKTLNPKNNEFFYR 73
69 IGKSD-SVTISVNNHKKHKQAGFLGCVR-----LISNAIN 105
74 VNPNSHRLLFVYDENRRLTRDD---FLGOVDP.LSHLPTEDPTMERPYTFKDFLLRPSH 130
106 RLKDTGYORLDCKLGPNDNDTVRGQIVVLSQSRDILGTGGVDCSRLLFNDN----- 158
139 KSRVKGFLRLKMAVFKNG-----GODENSDQRDMDHEGWEEVD-----SNDASQHOE 180
131 SNINCEELGP-----LPFGWEIN 307
159 -----LPDGMERRRTASGRIOYLNHITRTTOWERPT----- 189
161 ELPPPLPFGMEBKVDNLGRITVYVNNHNRRTQWHRP.SLMDVSSDNNIRQINQEAHRR 240
190 -----RPASEYSSP-----GRPLSCFVDENPTISGNT-----GATCGQSSDPR-- 227
241 FRSRHHISDLEBPSEGGDVPEPMEWT-ISEEVNIAGDSLGLALPPPPASGSRTPSQEL 299
228 -----LAER-----RVRS-----QRHNNYSRTHLHTPPDL 253
300 SEELSRRLQITTPDSNGEQFSSLIQREPSRLRSCSVTDAVAEQGHLPPPSVAVHTTPEL 359
254 PEGYEQRTTOOGQVYFLHTQGVSTWHD-----RVPRDL 288
368 PSGWEEBKDAKGRITVYVNNHNRRTTWTPTMQLAEDGASGATSNNNHLIEPQIRRPRL 419
289 SNINCEELGP-----LPFGWEIN 307
420 SSGTVTISAPLEGAKDSVPRRAVKDTLSNPQSPQSPYNSPKPOHKYTGSLPPGWEKRI 479

```

```

Oy 308 TATGRTVYFVDHNNRTTQFTDPRLSANLHLVLRNOLKDOQQQVSLCPD----- 359
Db 480 APNGRPPEIDHNTKITTWEDPRLKFVPM-----SKSLINPNDIGPLPFGW 526
Oy 360 -----TECLTVPR--YKRDLYQKXKILROELISQOQP 389
Db 527 EERHLDRITFYIHNSKITWEDPRLONPATIPAPVPSREFQKIDYRKKLKEAD 586
Oy 390 AGHCRIEVSREIEESYRQVKKRPPDLMK-RIMIKFGEGLDYGVAREMVLISH 448
Db 587 PNREMKLHRNNIEEESYRIRMSYKRPDVKARIMIFESBKIGDYGVAREMFLISKE 646
Oy 449 MLNRYGGLFOYSRDITLQINPSAY-NPEHLSYFHVGRIMGMAVFEHGYIDGGFTLP 507
Db 647 MKNRYGGLFEYSADNTLQINPSGLCNEDHLSYFTFGVAGLAFHFKLIDGGFTIR 706
Oy 508 FYKOLGKSTLIDMELVDPDLHNSLVLLENDITGVLDHPTCEVHNAYGEIIOHEIKPN 567
Db 707 FYKMLGKQITLNDMESVDSYNSLKMILENDPT-ELDLMFCIDENFGQYQVLDKPN 765
Oy 568 GKSIFVNEKKEVYRLVYNNRFLRGIEAQLAKQKFNVEVYQHLKTDEKLELIIC 627
Db 766 GSEIMVTNENREYIDVIOHMFVNRVQKQNNALFEGFTLLPDLIKIFDENELBLMC 825
Oy 628 GLGKIDVDMKVNTRLK-HCTPDSNIVKMWKAVFEFDEERARLLQFTGSSRPVLOG 685
Db 826 GLGVVDVDMRGHISYKNGYC-PHNPVIOHMFVNRVQKQNNALFEGFTLLPDLIKIFDENELBLMC 884
Oy 686 FKALQGAAPFLFTIHQIDACTNNLPKATCFNRIDIPYSEYEKYLITAEETCGF 745
Db 885 FAEIYGSNGPOLFTIEQWS-PEKLPRAHFCFNRIDIPYSEYEKYLITAEETCGF 943

```

RESULT 10

```

OQND8      PRELIMINARY;      PRT;      295 AA.
AC Q8ND8;
DT 01-OCT-2002 (TREMBLrel. 22, Created)
DT 01-OCT-2002 (TREMBLrel. 22, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE Hypothetical protein (Fragment).
GN DKF2564H23.3
OS Homo sapiens (human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
OX NCBI_TaxID=9606;
RN 1;
RP TISSUE=Brain.
RC SEQUENCE FROM N.A.
RA Wanduc R., Heubner D., Mewes H.W., Gassenhuber J., Wiemann S.;
RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL834242; CAD38919.1;
DR GO; GO:0005622; C:intracellular; IEA.
DR GO; GO:0004842; F:ubiquitin-protein ligase activity; IEA.
DR GO; GO:0005512; P:ubiquitin cycle; IEA.
DR InterPro; IPR000569; HECT_domain.
DR Pfam; PF00632; HECT_1.
DR SMART; SM00119; HECTC; 1.
DR PROSITE; PS50237; HECT; 1.
KM Hypothetical protein.
FT NON TER
SQ SEQUENCE 295 AA; 34030 MW; 0507325127A943EA CRC64;

```

```

Query Match 34.2%; Score 1379; DB 4; Length 295;
Best Local Similarity 85.4%; Pred. No. 9e-100;
Matches 251; Conservative 24; Mismatches 19; Indels 0; Gaps 0;
Oy 455 GLFOYSDDIYTLQINPSAVNPEHLSYFHVGRIMGMAVFEHGYIDGGFTLPFYKOLG 514
Db 2 GLFOYSTDNIMQINPDSSINPHLSYFHVGRIMGMAVFEHGYIDGGFTLPFYKOLG 61
Oy 515 KSLTLDMEVLDPDLHNSLVLLENDITGVLDHPTCEVHNAYGEIIOHEIKPKNSI PVN 574

```

```

Db 62 KPIQLSDLESVDPELHKLWLLENDITPVLDHPTCEVHNAYGEIIOHEIKPNRNPVT 121
Oy 575 EENKKEYYRLVYNNRFLRGIEAQLAKQKFNVEVYQHLKTPEKLELIICGAKIDV 634
Db 122 EENKKEYYRLVYNNRFLRGIEAQLAKQKFNVEVYQHLKTPEKLELIICGAKIDV 181
Oy 635 NDKNVTNLKCTPDSNIVKMWKAVFEFDEERARLLQFTGSSRPVLOGFKALQGAAG 694
Db 182 NDKNVTNLKCTPDSNIVKMWKAVFEFDEERARLLQFTGSSRPVLOGFKALQGAAG 241
Oy 695 PRLFTIHQIDACTNNLPKATCFNRIDIPYSEYEKYLITAEETCGF 748
Db 242 PRLFTIHQIDACTNNLPKATCFNRIDIPYSEYEKYLITAEETCGF 295

```

RESULT 11

```

OQ9CE3      PRELIMINARY;      PRT;      258 AA.
AC Q9CE3;
DT 01-JUN-2001 (TREMBLrel. 17, Created)
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE 2810411E22Rik protein (Fragment).
GN 2810411E22Rik.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN 1;
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Embryo;
RL MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamada S.,
RA Saito T., Okazaki Y., Gojopori T., Sono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA Schiml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barin G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamuya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombauts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whitaker C., Wilming L.,
RA Wyszewski-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohetsuki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
DR EMBL; AK013082; BAB28637.1;
DR MGI; MGI:1913563; 2810411E22Rik.
DR GO; GO:0005622; C:intracellular; IEA.
DR GO; GO:0004842; F:ubiquitin-protein ligase activity; IEA.
DR GO; GO:0005512; P:ubiquitin cycle; IEA.
DR InterPro; IPR000569; HECT_domain.
DR Pfam; PF00632; HECT_1.
DR SMART; SM00119; HECTC; 1.
DR PROSITE; PS50237; HECT; 1.
FT NON TER
SQ SEQUENCE 258 AA; 29670 MW; B0B804BC1FCB98 CRC64;

```

```

Query Match 34.1%; Score 1378; DB 11; Length 258;
Best Local Similarity 98.8%; Pred. No. 8.9e-100;
Matches 255; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
Oy 491 GMAVFEHGYIDGGFTLPFYKOLGKSTLIDMEVLDPDLHNSLVLLENDITGVLDHPTFC 550
Db 1 GMAVFEHGYIDGGFTLPFYKOLGKSTLIDMEVLDPDLHNSLVLLENDITGVLDHPTFC 60

```

```

QY 551 VERNHAYGELIQHLEKPKNGKSI PVNEENKKEVYRLVYNNRFLNGIEAOFLALOKGFNEYIP 610
DB 61 VERNHAYGELIQHLEKPKNGKSI PVTEENKKEVYRLVYNNRFLNGIEAOFLALOKGFNEYIP 120
QY 611 QHLLKTFDEKELELLICGLKIDVNDKMTNLKCTPDPSNIYKTFMKAVFDEERAR 670
DB 121 QHLLKTFDEKELELLICGLKIDVNDKMTNLKCTPDPSNIYKTFMKAVFDEERAR 180
QY 671 LLOFVYSSRVPLOGFKALOGAAGPRLFTIHQIDACTNNLPKRAHCFNRIDIPYSEYK 730
DB 181 LLOFVYSSRVPLOGFKALOGAAGPRLFTIHQIDACTNNLPKRAHCFNRIDIPYSEYK 240
QY 731 LYEKLTALIEETCGFAVE 748
DB 241 LYEKLTALIEETCGFAVE 258

RESULT 12
Q725N3 PRELIMINARY; PRT; 975 AA.
ID Q725N3;
AC Q725N3;
DT 01-OCT-2003 (TREMblrel. 25, Created)
DT 01-OCT-2003 (TREMblrel. 25, Last sequence update)
DT 01-OCT-2003 (TREMblrel. 25, Last annotation update)
DE NEDD4.2
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxId=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Maibert-Colas L., Nicolas G., Galand C., Lecomte M.-C., Dherry D.;
RT Identification of new partners of the epithelial sodium channel alpha
RT subunit."
RL Submitted (JUN-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY312514; AAF75706.1; -.
SQ SEQUENCE 975 AA; 111930 MW; AFD52AD504587B8 CRC64;

```

```

Query Match 34.0%; Score 1371; DB 4; Length 975;
Best Local Similarity 33.3%; Pred. No. 2, 1e-98;
Matches 332; Conservative 121; Mismatches 233; Indels 310; Gaps 26;

```

```

QY 14 LRLTVLCAKLVKDFRLPDPFAKV---VVDGSGQCH--STDVYKNTLDPMKQHYDLY 68
DB 22 LRVKVVSGIDLAKDIFGASDPYVLSIVYADENRRLALVQTKIKTNPKMNEEFYFR 81
QY 69 IGKSD-SVTISVNNHKKIKHKGAGFLGCYR-----LUSNAIN 105
DB 82 VNSPNHRLLEFVFDENRRLTRDD--FLGQVDPVLSHLPEDPTMERPYTFKDFLLRPRSH 138
QY 106 RLKDTGYORLDLCKLGNNDNTVAGQIVVSLQSRDRICTGGGVVDCSRFLPDND----- 158
DB 139 KSRVKGFLRLKAMYPNG-----GODENSDORDDMHGWVVD-----SNDASQHQK 188
QY 159 -----LPGWBERRTASGRIOYLNHTTTTQWERPT----- 189
DB 189 ELPPPLPPLPGMEERKVDNLGRITVYNNHNNRTTQWHRPSLMDVSSDNNIRQINQEAHRR 248
QY 190 -----PRASE-----YSSPG----- 199
DB 249 FRSRRLISEDLPEPPEGGGVPEPEWETISEEVNIAGDSLGLALPPPPASGKTSQELUS 308
QY 200 -----RPLSCFV----- 206
DB 309 EELSRRLQITPDNGEQLIOREPESRLRSCSVTDVAEQLLPPPSAPAGRASSTY 368
QY 207 -----DENP-----ISGTMGATCG----- 221
DB 369 TGGEEPTPSVAAYHTTGLPSGWEERKDAKGRITVYNNHNNRTTTRPIWQLAEDGASG 428
QY 222 -QSGDPLAEARRVVSQSHRYMSRT----- 245
DB 429 ATNSNNHLIPQIR--RPRLSSTVTLVLSAPLSEGAKDSPVRAVXDTLSNPSQSPSPYN 486

```

```

QY 246 -----HLHTPDDLEGEYEOERTTOGQVYFLHATQVSTWHDPRVP-----RDISNINCEE 295
DB 487 SPKQHKVQSFPLPGMEMELAPNGRPFFIDHTKTTTWEDPRLKPFVHRKSTKTSINPD 546
QY 296 LGRPLPGMEERILHLDGTFYIDHNSKITQWEDPRL-----QV----- 583
DB 547 LGRPLPGMEERILHLDGTFYIDHNSKITQWEDPRL-----QV----- 583
QY 356 CPDDTECLTVPR--YKRDLYQKILNQELSOQOPAGHCRIEVSREEIFEESRYQWK 413
DB 584 -----PAITGAVPVYSHEFFQKQVYFPRKDKKPADINFRFEMKLRHNNIFEESYRIMSV 638
QY 414 RPKDLAK-RMLTKRGEGLDYGVNAEMLYLSHEMLNRYTGLFQYSRDDITTLQINPD 472
DB 639 KRPVTLKARLWIEPESKGLDYGVAHEWFFLSKEMFNRYGLFEYSATDNTYLLQINPN 698
QY 473 SAV-NPEHLSYFHFVGRIMGAVPHGHYIDGFTLPPYKOLLGKSIITLDMELVDPDLN 531
DB 699 SGLCNEDHLSYFTTIGVAGLAVFHGKLDGFTFRPYKMLKQKITLNDMESVDSRYN 758
QY 532 SLVWILENDITGVLDHTFCVEHNAVGEIIQHEKPKNGKSI PVNEENKKEVYRLVYNNRFL 591
DB 759 SLKWIENDPT-ELDLNFCIDENNFQOTYQVLDKFNSEIMVTNENRREYIDVITQMRV 817
QY 592 RGIHAQFLALOKGFNEYIPQHLLKTFDEKELELLICGLKIDVNDKMTNLKCTPDPSNIYKTFMKAVFDEERAR 670
DB 818 NRVOQMAFLFEGFTLPLDLIKIPDENBELNCGLDVADVNDWQHSYKNGYIC-PN 876
QY 650 SNIVKMFKAVEFPDEERRALLOFVTGSSRVPLOGFKALOGAAGPRLFTIHQIDACTNN 709
DB 877 HPVIOFWKAVLMDAERIRLQFTVTSRVPKMAELIYSGNPQLFTEQWS--PEK 935
QY 710 LPKAHTCFNRIDIPYSEYKLYEKLTALIEETCGF 745
DB 936 LPKAHTCFNRIDIPYSEYKLYEKLTALIEETCGF 971

```

RESULT 13

```

Q725P2 PRELIMINARY; PRT; 967 AA.
ID Q725P2;
AC Q725P2;
DT 01-OCT-2003 (TREMblrel. 25, Created)
DT 01-OCT-2003 (TREMblrel. 25, Last sequence update)
DT 01-OCT-2003 (TREMblrel. 25, Last annotation update)
DE Ubiquitin ligase NEDD4g.
GN NEDD4L.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxId=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Qi H., Labrie C.;
RT "New Splicing Isoform of human Nedd4-2."
RT Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY112984; AAM76729.1; -.
KW L1gase.
SQ SEQUENCE 967 AA; 111339 MW; 86940A7588053957 CRC64;

```

```

Query Match 33.9%; Score 1369; DB 4; Length 967;
Best Local Similarity 33.3%; Pred. No. 3e-98;
Matches 332; Conservative 121; Mismatches 233; Indels 310; Gaps 26;
QY 14 LRLTVLCAKLVKDFRLPDPFAKV---VVDGSGQCH--STDVYKNTLDPMKQHYDLY 68
DB 14 LRVKVVSGIDLAKDIFGASDPYVLSIVYADENRRLALVQTKIKTNPKMNEEFYFR 73
QY 69 IGKSD-SVTISVNNHKKIKHKGAGFLGCYR-----LUSNAIN 105
DB 74 VNSPNHRLLEFVFDENRRLTRDD--FLGQVDPVLSHLPEDPTMERPYTFKDFLLRPRSH 130
QY 106 RLKDTGYORLDLCKLGNNDNTVAGQIVVSLQSRDRICTGGGVVDCSRFLPDND----- 158

```



```

131 KSRVKGFLRLKGMAYPKNG-----GQDENSDQDDMEHGEVND-----SNDASQHOZ 180
159 -----LPDGEERTASGRIOYLNHTRTTOWMEPT----- 189
181 ELPPPLPPGWEKVDNIGRTYYVNNHNRRTQWHRPSIMDVSSDNNIRQINQEAHRR 240
190 -----RPASE-----YSSPG----- 199
241 FRSRRHISEDLEPPESGEGVPEEWETISEEVNINAGSLGLALPPPPASGSGRTSPQELS 300
200 -----RPLSCFV----- 206
301 EELSRRLQITPDSNGEGFSSLIQREPSRSLSCSVTDAVAEQHLPSPAPAGRASSTV 360
207 -----IGTNGATG----- 221
361 TGESEPTSAVAVHTTLPGLSGWEERKADKGRYYVNNHNRRTTTPRIQOLAEDGASGS 420
222 -QSSDPRLAERRVNSQRHRYMSRT----- 245
421 ATNENNNHLEPQIR--RPRSLSPVTLASAPLEGASDPSVRAVKDTLSNPQSPQSPYV 478
246 -----HLTPPDLPEGIEQRTTQOGQVYFLATQTGVSTWHDPRVP-----RDLSNINCEE 295
479 SPKQHKKTQGLPPGWEEMRIAPNGRPFFIDHNTKITTWEDPRLKFPVHNSKTSLNEND 538
296 LGPLPGWEIRNTATGRVYFVDHNNRTTQFTDPLSANLHLVLRQQLDQOQOQVSVL 355
539 LGPLPGWEIRLHDGRFTFYIDHNSKITQWEDPVL-----QV----- 575
356 CPDDECTLPVR--YKDLVOKLILROELSQOQPOAGHCRIEVSREEIFEESYRQVYMK 413
576 -----PAITGPAVYSEBEFKQYDFPKKPKKPADINREEMKLRNNIPEESYRRIMSV 630
414 RPKDLWK-RLMIKRSGEGLDYGVAREMLYLSHEMLNPPYGLFOYSRDIYTLQINPD 472
631 KRPDLKRLWITBESEKGLDYGVAREMFFLSKEMFNPPYGLFELISADNTYTLQINPV 690
473 SAV-NPEHLISFHEFVGIMGMVPHGYIDGGFTLPPYKQLIGKSIITLDDMELVDDPLN 531
691 SGLCNEDHLSYFTIGVAGLAVPHGKLDGFFIRPYKMLGKQITLNDMESVDSYVYV 750
532 SLVWILNNDITGVLDHFCVFNHNYGELLIOHELKPNKSI-PVNEENKEVRYLYVMRPL 591
751 SLKWLILENDPT-ELDLMFCIDEENFGQYQVDLKNPSELMVTNENKREYIDLVQIMRV 809
592 RGIAPQFLALQKGENEVIPOHILKTFPEKELELIIIGLGRIDVNDWKVNTRLK--HCTPD 649
810 NRVCQKQWMAFLBEGTELLPDLIKIPDENELMLMGLGDVDVNDWQSHIYNQGYC-PR 868
650 SNIVKWKVAVEPFDERBARALLOFVTGSSRVPLQGFKALQGAAGPRLFTTHOIDACTNN 709
869 HPVLOQWPKAVLMDAKIRILLOFVGTGRVNPNGPAELVGSNGDPLFTIBOWGS-PRK 927
710 LPKAHTCFNRIDIPYSEYEKYLTLAIEETCGF 745
928 LPRAHCTCFNRIDLPYETFEELRKKLMAVENAGF 963

```

RESULT 14

```

Q9NT88 PRELIMINARY; PRT; 820 AA.
ID Q9NT88;
AC Q9NT88;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein (Fragment).
GN DKFZP434P2422.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Eumleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
NCBI_TaxID=9606;

```

```

RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Testis;
RA Blum H., Baueersachs S., Wewes H.W., Gassenhuber J., Wiemann S.,
RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: CONTAINS 1 C2 DOMAIN.
DR EMBL; AL137469; CAB70754.1; -.
DR PIR; T46412; T46412.
DR HSSP; Q33526; 1PTN.
DR GO; GO:0005622; C:intracellular; IEA.
DR GO; GO:0004842; F:ubiquitin-protein ligase activity; IEA.
DR GO; GO:0006512; P:ubiquitin cycle; IEA.
DR InterPro; IPR000008; C2.
DR InterPro; IPR008973; C2_CaB.
DR InterPro; IPR000569; HECT_domain.
DR InterPro; IPR002349; WW_Rsp5_WWP.
DR InterPro; IPR001202; WW_Rsp5_WWP.
DR Pfam; PF00168; C2_1.
DR Pfam; PF00397; WW_3.
DR Pfam; PF00632; HECT; 1.
DR PRINTS; PR00403; WMDOMAIN.
DR SMART; SM00239; C2_1.
DR SMART; SM00115; HECTC; 1.
DR SMART; SM00436; WW; 3.
DR PROSITE; PS50004; C2_DOMAIN_2; 1.
DR PROSITE; PS50237; HECT; 1.
DR PROSITE; PS01159; WW_DOMAIN_1; 3.
DR PROSITE; PS50020; WW_DOMAIN_2; 3.
KW Hypothetical protein.
FT NON_TER
SQ
SEQUENCE 820 AA; 95283 MW; 0FDB34B29B3F4123 CRC64;

Query Match 33.8%; Score 1364.5; DB 4; Length 820;
Best Local Similarity 37.1%; Pred. No. 5,4e-98;
Matches 315; Conservative 115; Mismatches 221; Indels 199; Gaps 25;

50 TDVYKTLDDPKKNQHYDLYKSD-SVITSVNKHKKHKGAGFLGCVR----- 98
12 TKTKIKTLNPKWNEEFYRVNPSNHRLLFEVFDENRLLTRD--FLGVDVPLSHLPTED 68
99 -----LISNAIRLKDTGYQRLDCKLGPNDNTVGRQIVVLSQSRDRTGTG 146
69 PTMERPTFFKDFLPRSHKSRVAGFLRKGMAYPKNG-----QDENSDQDDMEHGW 123
147 QVDCSRLLFND-----LPDGEERTASGRIOYLNHTRTTOWMEPT----- 189
124 EVVD-----SNDASQHOEBLPPPLPPGWEKVDNIGRTYYVNNHNRRTQWHRPSIMDV 178
190 -----RPAEYSSP-----GRPLSCFVDENTPISGTN-- 216
179 SSESNNIRQINQEAHRRFRSRHISEDLEPPESGEGVPEEWET-ISEEVNINAGDSLQ 237
217 -----GATCGQSSDPR-LAE--RVR-----SGRHRNYSRTHL----- 247
238 LALPPPPASGSGRTSPQELISEELSRRLQITPDSNGEQFSSLIQREPSRSLSCSVTDAVA 297
248 -----HTPDP----- 266
298 EOGHLRPEAKDSPVRAVKDTLSNPQSPQSPYVSPKQKVTGTFPLPGWEMRIAPNGR 357
267 VYFLHTQTGVSTWHDPRVP-----RDLSNINCEEGLPLPGWEIRNTATGGRVYFVDHNR 321
358 PFFIDHNTKTTTWEDPRLKFPVHNSKTSLNPNLGLPLPGWEMRIHLDGRFTFYIDHNSK 417
322 TTQFTDPLSANLHLVLRQQLDQOQOQVSVLCPDDECTLPVR--YKDLVOKLIL 379
418 ITQWEDPRL-----QV----- 449
380 ROELSQOQPOAGHCRIEVSREEIFEESYRQVYMKRPKDLWK-RLMIKRSGEGLDYGVVA 438
450 RKKLKRPADINREEMKLRNNIPEESYRRIMSVKRPDLVKARLWITFESEKGLDYGVVA 509
439 RWMYLLSHEMLNPPYGLFOYSRDIYTLQINPDPAV-NPBHLSYFHPVGRIGMAYVPHG 497

```


This Page Blank (uspio)


```

Db 59 SDSVTSVNNHKKIKKQAGFLGCVRLLSNAINRLKDTGYORDLCKLGPNDNDTVRGQ 118
QY 125 IVVSJQTRDRIGTGGVYDRCGLBNE-----GTVY----- 155
Db 119 IVVSJQSRDRIGTGGVYDRCGLBNE-----GTVY----- 178
QY 156 --EDGGPRPLSCFMEEPAPYDSTGAAGGNCRFVESPQDORLQARLNDVYGL 213
Db 179 ASYSPPGRLPSCFVDEMTPLSGINGATCG-----QSDPLAERRRQRSHNYM 229
QY 214 QTPONRPHGOSPELPEGEYQRTTVQGVYFLHTQTVSTWMDPRIPRLDINSVNCDELGP 273
Db 230 ---SRTHLHTPPDLPEGEYQRTTVQGVYFLHTQTVSTWMDPRIPRLDINSVNCDELGP 285
QY 274 LPPGHEVSVTSVGRIFYFDHNNRTTQFTDPR---LHIMNHQCCLEKPSQPLPLPSGSG 329
Db 286 LPPGHEIRNTATGRYFVDHNNRTTQFTDPRISANLHVLNRQOLKDOOQOQVY---S 341
QY 330 L---EDELPAQRYERDVLQKLVLRHLSLQOPAGHCRLEVSREEIFEESYQIMKMR 386
Db 342 LCPDTECLTVPRYRDVLQKLVLRHLSLQOPAGHCRLEVSREEIFEESYQIMKMR 401
QY 387 PKDLKRLMKVFRGEEGLDYGVAREMLYLLCHEMLNPPYGLFOYSTDNITMLOINPDSS 446
Db 402 PKDLKRLMKVFRGEEGLDYGVAREMLYLLCHEMLNPPYGLFOYSTDDIYTLQINPDSSA 461
QY 447 INPDLSYFHFYGRIMGAVFHGHYINGFTVPFYKOLLGKPIQLSDLESVDPRLHKSIV 506
Db 462 VNPBLSYFHFYGRIMGAVFHGHYIDGFTLPRFKOLLGKSTLDDMELVDPDLHNSLV 521
QY 507 WLENDITPVLDHTECVENHNAFGRILQHELKPNRGNVPTENKKEYRLLVYNNRFRGI 566
Db 522 WLENDITGVLDHTECVENHNAVGEIILQHELKPNKSGIFVNEENKKEYRLLVYNNRFRGI 581
QY 567 EAOFLALQKGFNEILPQHLKRPDOKLELITIGLDKIDLMDKSNRLKGCVADSNIVR 626
Db 582 EAOFLALQKGFNEVLPQHLKTFBKELELITIGLGKIDVMDKVNRLKHCPTDPSIVK 641
QY 627 WFMQAVETFEDEERRARLLQFTGSTRVPLQGFKALQSGTGAAGRLFTIHLIDANTNLP 686
Db 642 WFMKAVEFEDEERRARLLQFTGSSRVPLQGFKALQ---GAAGRLFTIHQIDACTNNLP 696
QY 687 KAHTCFNRIDIPYESYKLYEKLITAVEETCGFAVE 723
Db 699 KAHTCFNRIDIPYESYKLYEKLITAVEETCGFAVE 735

RESULT 2
US-09-392-163A-2
; Sequence 2, Application US/09392163A
; Patent No. 6503742
; GENERAL INFORMATION:
; APPLICANT: Beach, David H.
; APPLICANT: Calliguri, Maureen
; APPLICANT: Nefsky, Bradley
; TITLE OF INVENTION: Ubiquitin ligases, and Uses Related Thereto
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FOLEY, HOAG & ELIOT LLP
; STREET: One Post Office Square
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109-2170
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/392,163A
; FILING DATE:
; PRIOR APPLICATION DATA:

```

```

; APPLICATION NUMBER: US 08/539,205
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Vincent, Matthew P.
; REGISTRATION NUMBER: 36,709
; REFERENCE/DOCKET NUMBER: CSV-005.01
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 832-1000
; TELEFAX: (617) 832-7000
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 735 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-392-163A-2

Query Match 75.0%; Score 2913.5; DB 4; Length 735;
Best Local Similarity 72.9%; Pred. No. 7.1e-281;
Matches 552; Conservative 63; Mismatches 71; Indels 71; Gaps 9;

QY 5 IKRLTVLCAKLLAKKDFRLPDPFAKIVYDGGGCHSTDTYKNTLDPKRNQHYDLYGK 64
Db 12 VKRLTVL-----GLPDPFAKIVYDGGGCHSTDTYKNTLDPKRNQHYDLYGK 58
QY 65 TDSITSVNNHKKIKKQAGFLGCVRLLSNAISRLKDTGYORDLCKLNPSTDAVYRGQ 124
Db 59 SDSVTSVNNHKKIKKQAGFLGCVRLLSNAINRLKDTGYORDLCKLGPNDNDTVRGQ 118
QY 125 IVVSJQTRDRIGTGGVYDRCGLBNE-----GTVY----- 155
Db 119 IVVSJQSRDRIGTGGVYDRCGLBNE-----GTVY----- 178
QY 156 --EDGGPRPLSCFMEEPAPYDSTGAAGGNCRFVESPQDORLQARLNDVYGL 213
Db 179 ASYSPPGRLPSCFVDEMTPLSGINGATCG-----QSDPLAERRRQRSHNYM 229
QY 214 QTPONRPHGOSPELPEGEYQRTTVQGVYFLHTQTVSTWMDPRIPRLDINSVNCDELGP 273
Db 230 ---SRTHLHTPPDLPEGEYQRTTVQGVYFLHTQTVSTWMDPRIPRLDINSVNCDELGP 285
QY 274 LPPGHEVSVTSVGRIFYFDHNNRTTQFTDPR---LHIMNHQCCLEKPSQPLPLPSGSG 329
Db 286 LPPGHEIRNTATGRYFVDHNNRTTQFTDPRISANLHVLNRQOLKDOOQOQVY---S 341
QY 330 L---EDELPAQRYERDVLQKLVLRHLSLQOPAGHCRLEVSREEIFEESYQIMKMR 386
Db 342 LCPDTECLTVPRYRDVLQKLVLRHLSLQOPAGHCRLEVSREEIFEESYQIMKMR 401
QY 387 PKDLKRLMKVFRGEEGLDYGVAREMLYLLCHEMLNPPYGLFOYSTDNITMLOINPDSS 446
Db 402 PKDLKRLMKVFRGEEGLDYGVAREMLYLLCHEMLNPPYGLFOYSTDDIYTLQINPDSSA 461
QY 447 INPDLSYFHFYGRIMGAVFHGHYINGFTVPFYKOLLGKPIQLSDLESVDPRLHKSIV 506
Db 462 VNPBLSYFHFYGRIMGAVFHGHYIDGFTLPRFKOLLGKSTLDDMELVDPDLHNSLV 521
QY 507 WLENDITPVLDHTECVENHNAFGRILQHELKPNRGNVPTENKKEYRLLVYNNRFRGI 566
Db 522 WLENDITGVLDHTECVENHNAVGEIILQHELKPNKSGIFVNEENKKEYRLLVYNNRFRGI 581
QY 567 EAOFLALQKGFNEILPQHLKRPDOKLELITIGLDKIDLMDKSNRLKGCVADSNIVR 626
Db 582 EAOFLALQKGFNEVLPQHLKTFBKELELITIGLGKIDVMDKVNRLKHCPTDPSIVK 641
QY 627 WFMQAVETFEDEERRARLLQFTGSTRVPLQGFKALQSGTGAAGRLFTIHLIDANTNLP 686
Db 642 WFMKAVEFEDEERRARLLQFTGSSRVPLQGFKALQ---GAAGRLFTIHQIDACTNNLP 696
QY 687 KAHTCFNRIDIPYESYKLYEKLITAVEETCGFAVE 723
Db 699 KAHTCFNRIDIPYESYKLYEKLITAVEETCGFAVE 735

```

RESULT 3
US-08-539-205A-4
Sequence 4, Application US/08539205A
Patent No. 6001619
GENERAL INFORMATION:
APPLICANT: Beach, David H.
APPLICANT: Caligiuri, Maureen
APPLICANT: Nelsky, Bradley
TITLE OF INVENTION: Ubiquitin Ligases, and Uses Related Thereto
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: FOLEY, HOAG & ELIOT LLP
STREET: One Post Office Square
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109-2170
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/539,205A
FILING DATE: 04-OCT-1995
ATTORNEY/AGENT INFORMATION:
NAME: Vincent, Matthew P.
REGISTRATION NUMBER: 36,709
REFERENCE/DOCKET NUMBER: CSV-005.01
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 832-1000
TELEFAX: (617) 832-7000
INFORMATION FOR SEQ. ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 766 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-539-205A-4

Query Match 37.3%; Score 1471; DB 3; Length 766;
Best Local Similarity 40.9%; Pred. No. 4.2e-137;
Matches 327; Conservative 118; Mismatches 235; Indels 120; Gaps 20;
QY 4 SIKRLTLVLCAGKMLAKDPFLPDPFAKIVDGGGCHSDTVKNTLDPKMNCHYDLVYG 63
DB 7 SRIRITIVAADGLYKRDVFRFPDPRAVLTVDSG-QHTTTAKKTLNIPYMETFEVANT 65
QY 64 KTDSTITSVMNKKIKHKQAGFLGCVRLISNAISRDKDTGYRL--DLCKLNPSCTDAV 121
DB 66 DNSTIATIQVDDOKKF-KKKGQGLGVNLIVGDLVLDLAIIGDDMLTRDLKSNV--ENTVV 122
QY 122 RQGIIVSLQ-----TPDRIGGGSVVCORGLNEGTVYED 157
DB 123 HGKIIILSTTAOSTLOVPSSAASGARTORTSTINDPQSSKSSVSBNPASASGPTBD 182
QY 158 SGPG-----PPLSCFMEB---PAPV---TDSITG-----AAA 182
DB 183 NAPASPASSEPRFESFEDQYGLPFGWERRIDNLGRYYVDNHRSTTWIRPNLSVA 242
QY 183 GGGNCRVESP-----QDQRLAQRLNPDVYAGSLQTPONRPHGHSPLPESGYQ 234
DB 243 GAALAEHLSSASSANVTEGVQPSSSNARTEASVLTSTNATAG-----SGELPGEWBJ 286
QY 235 RTTYGQGVFLHTQTVGSTVHDPRIIPDLNSVN-----CDLGLPLPGWEVRS 283
DB 297 RYTBEGRPYFDVNDHTTRTTWVDPRRQYIISYSGPNNAITQQQPVSGGLPBGWEMRLT 356
QY 284 VSGRIYVDHNHNTQTQTPDRLHIMNHQCLKEPSQPLPLPBGSLDEDELPAQXYERD 343
DB 357 NTAAVYFVDHNTKTTTWDPR-----LPS--SL-DQNVF--QYKRD 392

QY 344 LVOKLTVLRHLSIQCPAGCHREIVSRREIPESRYQIMKMRKDLKRLMYKFRGEEG 403
DB 393 FRRKLIYFLSQPAL-HPLPGQCHIKVRNHI FEDSTAEIVRQSTADLKKLTKFDGEQ 451
QY 404 LDYGVAREWLYLICHEMLNPYGLFPQYSTDNITVMOQINPDSSINPDHLSYFHEVRIMG 463
DB 452 LDYGLSREYFPLSHMENPFCYCLFEYSVDNTYTLQINPHSGINPHELNFKFIGNVIG 511
QY 464 LAVFHGHYNGFVPPYKQILGKPIQLSDLESVDPLHLSIVLINDTLTPVDHNECV 523
DB 512 LAIFHRFVDAFPVVSFEYKQILQKRVTLQWESMDAIVSYSLVILINDTLTGVDLFSV 571
QY 524 EHNAFRLIQHELKPNRNVPVTEENKEVRYLYVMRPMRGIENAOPLAQGFENLIPQ 583
DB 572 EDNCFGEVVTIDLPKRNINTEVTEENKREYVDLVTV--IQRIIEQENAFHEGSELIPQ 630
QY 584 HLKPFQKELIELIIGLIDKIDLNDKSNTRLKCVAQDSNIVKRFQAVETPDEERARL 643
DB 631 ELINVFDERELELLIGSISBIDMEDMKKXDYRSSENDQIKFPMELMDEWSNEKKSRL 690
QY 644 LOFTGSTRVPLQGFEXLQSGTGAAPRLFTLIDANTDNLPRKHTCFNRIDIPYESSY 703
DB 691 LQFTTIGSRITVNGFKDQSSD--GPRKTIEXA-GEPMKLPKATHCPRDLDPPTSK 746
QY 704 EKYEKLTAVETCGFAVE 723
DB 747 KDLHKLSIAVEETIGFQGE 766

RESULT 4
US-09-392-163A-4
Sequence 4, Application US/09392163A
Patent No. 6503742
GENERAL INFORMATION:
APPLICANT: Beach, David H.
APPLICANT: Caligiuri, Maureen
APPLICANT: Nelsky, Bradley
TITLE OF INVENTION: Ubiquitin Ligases, and Uses Related Thereto
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: FOLEY, HOAG & ELIOT LLP
STREET: One Post Office Square
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109-2170
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/392,163A
FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/539,205
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Vincent, Matthew P.
REGISTRATION NUMBER: 36,709
REFERENCE/DOCKET NUMBER: CSV-005.01
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 832-1000
TELEFAX: (617) 832-7000
INFORMATION FOR SEQ. ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 766 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-392-163A-4
Query Match 37.9%; Score 1471; DB 4; Length 766;

Best Local Similarity 40.9%; Pred. No. 4.2e-137;
Matches 327; Conservative 118; Mismatches 235; Indels 120; Gaps 20;

[illegible]

RESULT 5
US-09-070-060-4
Sequence 4, Application US/09070060
Patent No. 5976849
GENERAL INFORMATION:
APPLICANT: Hustad, Carolyn M.
TITLE OF INVENTION: Human E3 Ubiquitin Proteinase
TITLE OF INVENTION: Lysase
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:
ADDRESSEE: ZENECA Pharmaceuticals, Inc.
STREET: 1800 Concord Pike

```

1 CITY: Wilmington
2 STATE: DE
3
4 COUNTRY: USA
5 ZIP: 19850-5437
6
7 COMPUTER READABLE FORM:
8 MEDIUM TYPE: Diskette
9 COMPUTER: IBM Compatible
10 OPERATING SYSTEM: DOS
11 SOFTWARE: FASTSEQ for Windows Version 2.0
12 CURRENT APPLICATION DATA:
13 APPLICATION NUMBER: US/09/070,060
14 FILING DATE: 30-APR-1998
15
16 CLASSIFICATION:
17
18 PRIOR APPLICATION DATA:
19 APPLICATION NUMBER: 60/073,839
20 FILING DATE: 05-FEB-1998
21
22 ATTORNEY/AGENT INFORMATION:
23 NAME: Higgins, Patrick H
24 REGISTRATION NUMBER: 39,709
25 REFERENCE/DOCKET NUMBER: PHM.70312
26 TELECOMMUNICATION INFORMATION:
27 TELEPHONE: 302.886.4889
28 TELEFAX: 302.886.8221
29
30 TELEX:
31
32 INFORMATION FOR SEQ ID NO: 4:
33 SEQUENCE CHARACTERISTICS:
34 LENGTH: 854 amino acids
35 TYPE: amino acid
36 STRANDEDNESS: unknown
37 TOPOLOGY: unknown
38
39 MOLECULE TYPE: peptide
40
41 US-09-070-060-4

```

Query Match	33.9%;	Score 1318;	DB 2;	Length 854;
Best Local Similarity	35.4%;	Pred. No. 8.9e-122;		
Matches 316;	Conservative 127;	Mismatches 229;	Indels 220;	Gaps 29;

[illegible]


```

QY 373 EIFESYRQIMKMKPKDLKRLMKVRFGEGLDYGVAREWLYLICHEMNPYYGLFOYS 432
DB 507 TLFEDSFQOIMSPSPDQLRRRLWVIFPGSEGLDYGVAREWFLLSHEVLANPWYCLPEYA 566
QY 433 TDNIYMLQINPDSSINPDHLSYFHFVGRIMGAVFGHYINGFTVPFYKQLGKPIQLS 492
DB 567 GKNVYCLQINPASYINPDHLYFRFGRIAMALFHGKFIIDTGFSLPFYKRLINKPVGLK 626
QY 493 DLESVDPELHKSJLVILENDITPV-LDHTFCVENHAGRILOHELKPRGNRPVTEENK 551
DB 627 DLESIDPEFYNLSLWKENNIEECGLMEYFSDYKELIGISHDLKPRGNRIIVTEENK 686
QY 552 EYRLVYVMRFMRGIEAFLAQKGFNELIPQHLKPFQKELIIGGLDKIDLNDWS 611
DB 687 EYIRVAVMRSLRGVEBQTAFFEGFENILPQYLYQYDALEVLICMGEBIDLNDWR 746
QY 612 NTRLKHCVADSNIVRFWQAVETPDEERRALLQPVYGTSTVPVLOGFKALOGSTGAAGR 671
DB 747 HAIYHYRTSKQIMFWQFVKEIDNEKRMRLQVGTGCLPVGGFADLNGSN--GPO 803
QY 672 LFTIHLIDANTDNLPKATCFNRIDIPYSESYEKLYEKLTAVEETCGFAVE 723
DB 804 KFCIEKV-GKENWLPKSHTCNRLDLPYKSYEOLKELFLAETETGEGOE 854

```

RESULT 6 US-09-357-746-4

```

; Sequence 4, Application US/09357746
; Patent No. 6087122
; GENERAL INFORMATION:
; APPLICANT: ZENECA Limited
; TITLE OF INVENTION: HUMAN E3 UBIQUITIN PROTEIN LIGASE
; FILE REFERENCE: PHN.70312.NI
; CURRENT APPLICATION NUMBER: US/09/357,746
; CURRENT FILING DATE: 1998-07-21
; EARLIER APPLICATION NUMBER: US No. 6087122 60/073,839
; EARLIER FILING DATE: 1998-02-05
; EARLIER APPLICATION NUMBER: US No. 608712209/070,060
; EARLIER FILING DATE: 1998-04-30
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 4
; LENGTH: 854
; TYPE: PRT
; ORGANISM: Mus musculus
; US-09-357-746-4

```

Query Match 33.9%; Score 1318; DB 3; Length 854;
Best Local Similarity 35.4%; Pred. No. 8.9e-122; Indels 220; Gaps 29;
Matches 316; Conservative 127; Mismatches 229;

```

QY 6 KIRLVYCAK-NLAKDQFRLDPFAKIVVDGSGCHSTDTVKNITLDPKMNQHYDIYVGK 64
DB 9 QLGITVSAKLEKMKKMPFG-PSPYVEVTD--GQSKTEKCNNTNPFKKQPLVIVTP 65
QY 65 TDSITIVYVNAHKIKKKQAGFLCCVLA-----LSNAISLKTGTGR----- 107
DB 66 TSKLCFRWVSHQTL--KSDVLTGTAGIDYITLKSNNMKLEEVVMTLQVGDKEPTEIM 122
QY 108 --LDLCKLNPEDTAVAGQIVS-----LQTRD--RIGTGS--VYDCR 145
DB 123 GDLVLC---LDGLOVEAEVYTNGETSCSESTQNDGCGTRDTRSTNGSEDEPEVAAS 178
QY 146 G-----LIENGTVEGSGPGRPLSCNEEPAPY-----DSTG 179
DB 179 GENFRANGNNSPSLSNG--FKPSRPPRP-----SRPPPTPRRPAVNGSPSTNSDSDG 231
QY 180 AA-----AGGNCRFVSPSQ-----DQRLQQR 203
DB 232 SSTGSLPPTNNVNTSTEGATSGILPLITSSGSGRPPLNTVQALPFGMEKRV----- 267
QY 204 LRNPVPRGSL-----QTPQNPFGHQSPELEGEYEQRTTVGGQVYFLHTGTGVSTWH 255

```

```

DB 288 ----DQGRVYVVDHVEKRTWDRP-----PLPFGMEKRVDNKRIYYVDHFRTITWQ 338
QY 256 DPRIP-----RL-----NSVNDELGPRLPGMEVR 281
DB 339 RPTLESVANYBQWOLQBSOLQAGAQGNORFIYGNQDLFATSOKEFPDPLGPPFGMEKR 398
QY 282 STVGRIYFVHNHNRITQFTDPRLHIMNHQCLKEPSQPLPL----- 324
DB 399 TDSNGRIVFVHNHNRITQWEDPR-----SQGLNE--KPLPBGMEKRVTVDGIPYFVDH 450
QY 325 -----PESGLEDEELPAQRYEDYQKLYR--HELSLOQPQAGCHGIEVSR 372
DB 451 NRRATYIDPRTGKSLDNGPQIAYVRDFKAKVQYFFRWCCQLAMPQ---HITVTRK 506
QY 373 EIFESYRQIMKMKPKDLKRLMKVRFGEGLDYGVAREWLYLICHEMNPYYGLFOYS 432
DB 507 TLFEDSFQOIMSPSPDQLRRRLWVIFPGSEGLDYGVAREWFLLSHEVLANPWYCLPEYA 566
QY 433 TDNIYMLQINPDSSINPDHLSYFHFVGRIMGAVFGHYINGFTVPFYKQLGKPIQLS 492
DB 567 GKNVYCLQINPASYINPDHLYFRFGRIAMALFHGKFIIDTGFSLPFYKRLINKPVGLK 626
QY 493 DLESVDPELHKSJLVILENDITPV-LDHTFCVENHAGRILOHELKPRGNRPVTEENK 551
DB 627 DLESIDPEFYNLSLWKENNIEECGLMEYFSDYKELIGISHDLKPRGNRIIVTEENK 686
QY 552 EYRLVYVMRFMRGIEAFLAQKGFNELIPQHLKPFQKELIIGGLDKIDLNDWS 611
DB 687 EYIRVAVMRSLRGVEBQTAFFEGFENILPQYLYQYDALEVLICMGEBIDLNDWR 746
QY 612 NTRLKHCVADSNIVRFWQAVETPDEERRALLQPVYGTSTVPVLOGFKALOGSTGAAGR 671
DB 747 HAIYHYRTSKQIMFWQFVKEIDNEKRMRLQVGTGCLPVGGFADLNGSN--GPO 803
QY 672 LFTIHLIDANTDNLPKATCFNRIDIPYSESYEKLYEKLTAVEETCGFAVE 723
DB 804 KFCIEKV-GKENWLPKSHTCNRLDLPYKSYEOLKELFLAETETGEGOE 854

```

RESULT 7

US-09-070-060-3
Sequence 3, Application US/09070060
Patent No. 5976849

```

; GENERAL INFORMATION:
; APPLICANT: Husted, Carolyn M.
; TITLE OF INVENTION: Human E3 Ubiquitin Protein
; TITLE OF INVENTION: Ligase
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ZENECA Pharmaceuticals, Inc.
; STREET: 1800 Concord Pike
; CITY: Wilmington
; STATE: DE
; COUNTRY: USA
; ZIP: 19850-5437
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/070,060
; FILING DATE: 30-APR-1998
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/073,839
; FILING DATE: 05-FEB-1998
; ATTORNEY/AGENT INFORMATION:
; NAME: Higgins, Patrick H.
; REGISTRATION NUMBER: 39,709
; REFERENCE/DOCKET NUMBER: PHN.70312
; TELECOMMUNICATION INFORMATION:

```

/ TELEPHONE: 302.886.4899
 / TELEFAX: 302.886.8221
 / TELEX:
 / INFORMATION FOR SEQ ID NO: 3:
 / SEQUENCE CHARACTERISTICS:
 / LENGTH: 852 amino acids
 / TYPE: amino acid
 / STRANDEDNESS: unknown
 / TOPOLOGY: unknown
 / MOLECULE TYPE: peptide
 / US-09-070-060-3

Query Match 33.5%; Score 1301; DB 2; Length 852;
 Best Local Similarity 34.7%; Pred. No. 4,4e-120;
 Matches 308; Conservative 126; Mismatches 241; Indels 212; Gaps 26;

```

QY 6 KIRLTVLCAK-NLAKKDFRLPDPFAKIVDVGSGCHSTDTVKNLTDPKMNQHYDLYVGK 64
DB 9 QLOITVISAALKENKKMWF--PSFYVEVTVD--GSKKTEKCNNTNSPKMKQPLVITVP 65
QY 65 TDSITISVNNKKIHKK--QGAGFLGCVRLSNLSRLKDT-----GYQR-----L 108
DB 66 VSKLHFRVWSHOTLSKSDVLLGTALDIYETLKSNNMKLEVVVTLQGGDKPETTIGDL 125
QY 109 DLCKLNPEDTDAVRGOIVVSIQT-----RDRICTGGS-----VYDC 144
DB 126 SIC-----LDGLQLEBEVVTNGETTSBSASQNDGSRKDETRVSTNGSDDEPDAGAGN 181
QY 145 RGL-----LENSTVYEDSGPGRPLSCFMEEPAPYTDSTGAAGGNCRFVES----- 192
DB 182 RRVSGNNSPSSLNGG--FKPSRPPRP-----SRPPPPRRRPAASVNGSPSATSESDGST 234
QY 193 -----PSODORLOAQRLNPD 208
DB 225 GSLPPTNTNTSSEATSGLIIPLTISGSGRPRLNPVTQAPLRPGMEQRY-----D 286
QY 209 VRGSL-----QTPQNRPHGHOSPELPEGYEQRTTVQGVYFLHTQGVSTWHDPRIP 260
DB 287 QHGRVYVVDHVEKRTTWDRPE-----PLPPGMRERVDNMGRIVYVDHFTRTTQWRPTLE 341
QY 261 -----RDL-----NSVNCDELGLPRGMEVRSYVSG 286
DB 342 SVRNYEQWQLQSGAQMGQFNOFRIFYGNODLFAITSOSKEFDPLGPLPQMEKRTDSNG 401
QY 287 RIYFVDHNNRTTOFTDPRLLHIMNHQCOLKEPSQPLP----- 324
DB 402 RYVFVNHNTRITQWMDPR-----SQGLNE--KPLPGMEMRFTVDSIPYVDHNRRTT 453
QY 325 -----PSGSLDEDELPAPORYEDLVQKLKYLK--HELSLQOPAGHCRLEVSSEEFEE 377
DB 454 TYIDRRTGKSALDNGPQIAYVDRFKAQYFRFMCQOLAMPQ-----HIKIVTRKTLFED 509
QY 378 SYROIIMKRPKDLKRLKLVKFRGEGLDYGGVAREMTYLLCHEMILNFPYGLFOYSTDNLY 437
DB 510 SFQQLMSSFPQDLRRLRVIFPGEGLDYGGVAREMFFLSHEVLNMPYCLFEYAGKDNV 569
QY 438 MLOINPDSSINPDHLSYHFVGRIMGLAVFPGHYINGFTVPFYKQLGKPIQLSDLESV 497
DB 570 CLOINPASYINPDHLYFRFGRIFAMALFPGKFIIDGFSLPFYKRIINLPVGLKDESI 629
QY 498 DPELHKSJLVILENDITPV--LDHFFCVENHAFGRILQHEHLEKPNRANVPVTEKKXEVRL 556
DB 630 DPEFNSILVYKKNIEBCDLEMTFSVDKELIGIKSHDLKPNQGNILVLEENKEEYIRK 669
QY 557 YVNWEMFGIAPLOKGFNELIPQHLKPFQDKLELIIIGLDKIDINDMKSNNRLK 616
DB 630 VAEWRLSKGVSEQQAPEEGEFNELIPQYLOQYFPAKLEVLVLCMQGIDINDMQRAIYR 749
QY 617 HCVADSNIVREKQAVETFEDEBARLLQFTVGTSTRPLQGFKALQSTGAAPRLFTIH 676
DB 750 HYAATSKQIMFWQFVKEIDNEKMRLLQFTGTGCRLPVGGFADLMKSN--GPQKFCIE 806
QY 677 LIDANTDNLPRKATCFNRIDIPYSEYKLYEKILLTAVEETCGPAVE 723

```

DB 807 KY-GKENWLPBSHTCFNRIDLPYKSYEQLEKELFAIEETGFGQ 852

RESULT 8
 US-09-357-746-3
 / Sequence 3, Application US/09357746
 / Patent No. 6087122
 / GENERAL INFORMATION:
 / APPLICANT: ZENEGA Limited
 / TITLE OF INVENTION: HUMAN E3 UBIQUITIN PROTEIN LIGASE
 / FILE REFERENCE: PHM 70312.N1
 / CURRENT APPLICATION NUMBER: US/09/357,746
 / CURRENT FILING DATE: 1999-07-21
 / EARLIER APPLICATION NUMBER: US No. 6087122 60/073,839
 / EARLIER FILING DATE: 1998-02-05
 / EARLIER APPLICATION NUMBER: US No. 608712209/070,060
 / EARLIER FILING DATE: 1998-04-30
 / NUMBER OF SEQ ID NOS: 15
 / SOFTWARE: FastSeq for Windows Version 3.0
 / SEQ ID NO 3
 / LENGTH: 852
 / TYPE: PRT
 / ORGANISM: Homo sapiens
 / US-09-357-746-3

Query Match 33.5%; Score 1301; DB 3; Length 852;
 Best Local Similarity 34.7%; Pred. No. 4,4e-120;
 Matches 308; Conservative 126; Mismatches 241; Indels 212; Gaps 26;

```

QY 6 KIRLTVLCAK-NLAKKDFRLPDPFAKIVDVGSGCHSTDTVKNLTDPKMNQHYDLYVGK 64
DB 9 QLOITVISAALKENKKMWF--PSFYVEVTVD--GSKKTEKCNNTNSPKMKQPLVITVP 65
QY 65 TDSITISVNNKKIHKK--QGAGFLGCVRLSNLSRLKDT-----GYQR-----L 108
DB 66 VSKLHFRVWSHOTLSKSDVLLGTALDIYETLKSNNMKLEVVVTLQGGDKPETTIGDL 125
QY 109 DLCKLNPSTDAVRGOIVVSIQT-----RDRICTGGS-----VYDC 144
DB 126 SIC-----LDGLQLEBEVVTNGETTSBSASQNDGSRKDETRVSTNGSDDEPDAGAGN 181
QY 145 RGL-----LENSTVYEDSGPGRPLSCFMEEPAPYTDSTGAAGGNCRFVES----- 192
DB 182 RRVSGNNSPSSLNGG--FKPSRPPRP-----SRPPPPRRRPAASVNGSPSATSESDGST 234
QY 193 -----PSODORLOAQRLNPD 208
DB 225 GSLPPTNTNTSSEATSGLIIPLTISGSGRPRLNPVTQAPLRPGMEQRY-----D 286
QY 209 VRGSL-----QTPQNRPHGHOSPELPEGYEQRTTVQGVYFLHTQGVSTWHDPRIP 260
DB 287 QHGRVYVVDHVEKRTTWDRPE-----PLPPGMRERVDNMGRIVYVDHFTRTTQWRPTLE 341
QY 261 -----RDL-----NSVNCDELGLPRGMEVRSYVSG 286
DB 342 SVRNYEQWQLQSGAQMGQFNOFRIFYGNODLFAITSOSKEFDPLGPLPQMEKRTDSNG 401
QY 287 RIYFVDHNNRTTOFTDPRLLHIMNHQCOLKEPSQPLP----- 324
DB 402 RYVFVNHNTRITQWMDPR-----SQGLNE--KPLPGMEMRFTVDSIPYVDHNRRTT 453
QY 325 -----PSGSLDEDELPAPORYEDLVQKLKYLK--HELSLQOPAGHCRLEVSSEEFEE 377
DB 454 TYIDRRTGKSALDNGPQIAYVDRFKAQYFRFMCQOLAMPQ-----HIKIVTRKTLFED 509
QY 378 SYROIIMKRPKDLKRLKLVKFRGEGLDYGGVAREMTYLLCHEMILNFPYGLFOYSTDNLY 437
DB 510 SFQQLMSSFPQDLRRLRVIFPGEGLDYGGVAREMFFLSHEVLNMPYCLFEYAGKDNV 569
QY 438 MLOINPDSSINPDHLSYHFVGRIMGLAVFPGHYINGFTVPFYKQLGKPIQLSDLESV 497
DB 570 CLOINPASYINPDHLYFRFGRIFAMALFPGKFIIDGFSLPFYKRIINLPVGLKDESI 629

```

QY 498 DELHAKSLWILNDITPV-LDHTCEVFNAGRILOHELRKNGRNPVTEENKEYVRL 556
 Db 630 DEEFVNSLIMVKNENNEECDDLEMYSDVDEILGEIKSHDKNGGAILVTEENKEYVRL 689
 QY 557 YVNMREAGIEAOFALQKGFNELLIPQHLKRPDCKELELIGLKDIDLNDMKNSTRBK 616
 Db 690 VAEWRLSRVGEBOQAFFEGFNEILLPOQYLOQFDAELEVLICGMQEIDLNDMKNHAIYR 749
 QY 617 HCVADSNIVRWQAVETPDEERRAPLLOFVGTSTVPLQGTALOGSTGAAGPRLFTTH 676
 Db 750 HVAETSKQIMFWQFKEIDNEKRMRLLOFVGTCTCLPVGFGADMGNSN---GPOKFCLE 806
 QY 677 LIDANTDNLPKATCFNRIDIPYSESEKYLEKLTAVEETCGFAVE 723
 Db 807 KY-GKENMLPRSHTCFNRIDLPPYKSEYEQKELFLAIBETGFGCE 852

RESULT 9

US-08-895-601-6
 ; Sequence 6, Application US/08895601
 ; Patent No. 6060263
 ; GENERAL INFORMATION:
 ; APPLICANT: Beer-Romero, Peggy
 ; APPLICANT: Strack, Peter J.
 ; APPLICANT: Glass, Susan J.
 ; APPLICANT: Rolfe, Mark
 ; TITLE OF INVENTION: REGULATION OF KAPPA B (1KB) DEGRADATION,
 ; TITLE OF INVENTION: AND METHODS AND REAGENTS RELATED THERETO
 ; NUMBER OF SEQUENCES: 16
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: FOLEY, HOAG & ELIOT LLP
 ; STREET: One Post Office Square
 ; CITY: Boston
 ; STATE: MA
 ; COUNTRY: USA
 ; ZIP: 02109-2170
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/895,601
 ; FILING DATE: 16-JUL-1997
 ; CLASSIFICATION: 435
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Vincent, Matthew P.
 ; REGISTRATION NUMBER: 36,709
 ; REFERENCE/DOCKET NUMBER: MITV-096.01
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 617-832-7000
 ; TELEFAX: 617-832-7000
 ; INFORMATION FOR SEQ ID NO: 6:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 927 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; US-08-895-601-6

Query Match 32.4%; Score 1259; DB 3; Length 927;
 Best Local Similarity 33.6%; Pred. No. 7, 8e-116;
 Matches 307; Conservative 129; Mismatches 239; Indels 240; Gaps 28;
 QY 7 IRLTVLCANLAKKDFRLDPPEAKIV---VDGSGCHSTDTVAKTLDPKKN----- 55
 Db 48 VRVAVIGAGIAGKDDIGASDPYRVTVLPDPMGVLTSLVQTKIKSLINKMNEELIFRY 107
 QY 56 ---OH-----YLVYKTD--SITISVNN-----HKIKKQ 83
 Db 108 HPOQHRLLEFVFNENRLLTRDDQVDPVLPYLPTEPRLERPYTFKDFVLRSHKRV 167

QY 84 AGFLGCVALLSNARLRLKDT-----GYQRDL-----LCKL----- 113
 Db 168 KGYL-----RLKMTYLPKTSSEDDNNAQAELEPQWVWVLDQPDAAHQOQCE 216
 QY 114 -----NPSD--TDVARGQIVSLQ-----TRDR 134
 Db 217 PSLPLPGMEERQDILGRTYYVNHESRRTQWRPPTPQNLTDENGNLOQAOAFTTRQ 276
 QY 135 IGTGGSVDCRGLLEN-----EGTYEDSG--PGRSLSCFMEEPAYTD----- 176
 Db 277 ISEETESVDNOSSENNWEIIRDEBATVYSSQAFSPSPSSNLDPYTHLABELNRLITFG 336
 QY 177 ---STGAAGGNCR-----FVSPS-----QDQIQAQRL-- 204
 Db 337 NSAVSOPASSNHSRRGSLQAYTFEQRPLPVLLPTSSGLPGMEERQDEGRSYVVDH 396
 QY 205 -----RNPVRSLSIQPQ-----NRPHGQSPLEPEGYEQ 234
 Db 397 NSRTTWTKPTVQAVETSQLTSQSSAQPOSQASTDSGOQVTPQSEIQQFLPKMEV 456
 QY 235 RTVQGVYFPLHTQGVSTWHDPR--IPRLINSV---NDELGRLPPGMEVRSTVSGRI 288
 Db 457 RHAPNGRPFIDHTKTTTWEDPRLKIPALHKGKSLDTSNDGLPLPGMEERTHTDGR 516
 QY 289 YFVDHNNRTQFTDPRLHIMNHOCOLKEBPQPLPFSGLDEBELPAQRYERDLYQKL 348
 Db 517 FYINNIRKTQWEDPRLENV---ATGPAPV-----YSRDYRKY 553
 QY 349 KYLRHESLSLOQPOAGHRIEVSREIFEESYQIOMKAPKD-LKKRLMVKFRGEGLDYG 407
 Db 554 EFRKRLKQNDINPKEMKARATVLEDSYRIMGKADFLARLMIERDEGKGLDYG 613
 QY 408 GVAREWLYLCEKLNIPYGLFOYSTDNITWMLQINPDSS-LNPHLSYFHVGVGIMGLAV 466
 Db 614 GVAREWFLISKEMNPYGLFEYSATDNYTLQINPNSGLCNEHLSYFKIGVAGAV 673
 QY 467 FHGHYNGFTVPFQKOLGKPIQSDLESVDPELHSLWILNDITPVLDHTCFVCEV 526
 Db 674 YHGKLDGFTFRPKMLKPIQLDMESVDSEYINSLRILINDPRL-IDLFTIIDE 732
 QY 527 AFGRILOHELRKNGRNPVTEENKEYVRLVYVNRFPNGIEAQLALQKGFNELLIPQHL 586
 Db 733 LFGQTHQHELRKNGSEIIVTNKKKEYIYLVQRFVNRIOKMAAFKGFELIPDL 792
 QY 587 KPFDQKELELIGLKDIDLNDMKNSTRBK-CVADSNIVRWQAVETPDEERRAPLLO 645
 Db 793 KIPFENELELMCGAGDVADMDREHTYKNGYSANHQVIOFWKAVLMDSEKIRLLQ 852
 QY 646 FVTGSTRVPLQGTALOGSTGAAGPRLFTTHLIDANTDNLPKATCFNRIDIPYSESEK 705
 Db 853 FVTGSTRVPMNGFALYGSN---GQOSTVYQW--GTPKRLRATHTCFRLDLPYSESE 908
 QY 706 IYKELTAVEETCGF 720
 Db 909 LMDKLQMAIENTQGF 923

RESULT 10

US-08-539-205A-6
 ; Sequence 6, Application US/08539205A
 ; Patent No. 6001619
 ; GENERAL INFORMATION:
 ; APPLICANT: Beach, David H.
 ; APPLICANT: Caligiuri, Maureen
 ; APPLICANT: Nefsky, Bradley
 ; TITLE OF INVENTION: Ubiquitin Ligases, and Uses Related Thereto
 ; NUMBER OF SEQUENCES: 6
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: FOLEY, HOAG & ELIOT LLP
 ; STREET: One Post Office Square
 ; CITY: Boston
 ; STATE: MA
 ; COUNTRY: USA

ZIP: 02109-2170
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/539,205A
FILING DATE: 04-OCT-1995
ATTORNEY/AGENT INFORMATION:
NAME: Vincent, Matthew P.
REGISTRATION NUMBER: 36,709
REFERENCE/DOCKET NUMBER: CSV-005.01
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 832-1000
TELEFAX: (617) 832-7000
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 834 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-539-205A-6

Query Match 31.6%; Score 1228.5; DB 3; Length 834;
Best Local Similarity 40.4%; Pred. No. 7.1e-113;
Matches 276; Conservative 103; Mismatches 211; Indels 93; Gaps 19;
93 LSNAISRL-----KDTGYQRLDCKLNPSD-----TDAVRGQ-----IYVSLQTRDR 134
186 LSEELSRRLQITPPSNGEQFSSLIQREPSRLRSCSVTDVAEAGHLPPPSVAVYHTTGP 245
135 IGTGSAV-VDRCGLL-----ENEGTVYEDSGRPLSCFMEEPAPYDSTGAAGAGN-- 186
246 LPSGWEERKDAKGRYYVYVNNHNRRTTWT-----RPIMQLABDGA-----SGSATNSNHL 295
187 -----CRFVESPQDQRLQARLRNPVYGSU-----QTPQNPHGQSPSE----- 227
226 IEPQIRRRSLSPFTVLISAPLEGAKOSPVRAYADTISNPQSPQSPYNSPKQHKYTK 355
228 --LPEGYEQRTTVQGVYFLHTQGVSTWHDRIIP-----RDLSVNCDELGRPLPPGWEV 280
356 SFLPPGWEMLRILAPNRPFFIDHNKRTTWEDPRILKFPVHMSKSLNPNNDLGPPLPGWEE 415
281 RSTVSGRIYFDNHNRTQFTDPRILHIMNHQCQKESQCLPLPSEGSLEDELPAPORY 340
416 RIHLDGRFTFYIDHNSKITQWEDPRILQ-----PALTG-----PAVPY 452
341 ERDLVQKILVLRHLSLQOPQAGHCRLEVSREEIFEESYRQIMKM-RPKDLKRLMYKFR 399
453 SRBFQOKYDYFRKRLKPRADIPNRFPEMKLHNNIFEESEYRILMSVKRPDLKARLWIEFE 512
400 GEEGLDVGVARREMLYLICHEMLNRYGLPQYSTDNLYMLOINPDSI--NPDHLSYHNFV 458
513 SEKGLDVGVARREMLFLSKEMENPYGLFEYSATDNTYLIQINNSGLCNEDHLSYTFI 572
459 GRINGLAVFHGYNGFTVFPYKQLLQKPIQLSDLESVDPELHKSJLWILENDITPVLD 518
573 GRVAGLAVFHGKLDGFEIRFPFYKMLGKQITLNDMESVSEYNSJLWILENDITE--LD 631
519 HTFCVENAFGRILIOHLEKPRGNRVPTENKKEKYVLLVNNRPMRILEQOFLALQKGRN 578
632 LMFCDIEENFQYQVVDIKRPGSEIYVNNKREYIDLVIQMRVNVVQKQNAFLLEGFT 651
579 ELIPQHLKPRPDQKLELILIGLIDKIDLNDKSNTRLKHCVADSN--IVRFMOAVERFDE 637
692 ELIPIDIKIFDENELMLMGGLDADVNDMRQHSIKNGYCPRHPIYQVFWAYVLMDA 751
638 ERARLLOFTVGSIRVPLQGFKALQSGTGAAGAPPLFIHLIDANTDNLPRALHCENRIDI 657
752 EKRIIRLOFTVGTGRVPMNGFAELVGSN--GPOFLTEIQM--GSPKDLPRALHCENRIDI 807
QY 698 PPYESYEKLYEKLTAVEETGCF 720

DB 808 PPYEFEDLREKLLMAVENAQGF 830

RESULT 11
US-09-392-163A-6
Sequence 6, Application US/09392163A
Patent No. 6503742
GENERAL INFORMATION:
APPLICANT: Beach, David H.
APPLICANT: Nefsky, Bradley
APPLICANT: Caligiuri, Maureen
TITLE OF INVENTION: Ubiquitin Ligases, and Uses Related Thereto
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: FOLEY, HOAG & ELIOT LLP
STREET: One Post Office Square
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109-2170
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/392,163A
FILING DATE:
PRIOR APPLICATION NUMBER:
APPLICATION NUMBER: US 08/539,205
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Vincent, Matthew P.
REGISTRATION NUMBER: 36,709
REFERENCE/DOCKET NUMBER: CSV-005.01
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 832-1000
TELEFAX: (617) 832-7000
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 834 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-392-163A-6

Query Match 31.6%; Score 1228.5; DB 4; Length 834;
Best Local Similarity 40.4%; Pred. No. 7.1e-113;
Matches 276; Conservative 103; Mismatches 211; Indels 93; Gaps 19;
93 LSNAISRL-----KDTGYQRLDCKLNPSD-----TDAVRGQ-----IYVSLQTRDR 134
186 LSEELSRRLQITPPSNGEQFSSLIQREPSRLRSCSVTDVAEAGHLPPPSVAVYHTTGP 245
135 IGTGSAV-VDRCGLL-----ENEGTVYEDSGRPLSCFMEEPAPYDSTGAAGAGN-- 186
246 LPSGWEERKDAKGRYYVYVNNHNRRTTWT-----RPIMQLABDGA-----SGSATNSNHL 295
187 -----CRFVESPQDQRLQARLRNPVYGSU-----QTPQNPHGQSPSE----- 227
226 IEPQIRRRSLSPFTVLISAPLEGAKOSPVRAYADTISNPQSPQSPYNSPKQHKYTK 355
228 --LPEGYEQRTTVQGVYFLHTQGVSTWHDRIIP-----RDLSVNCDELGRPLPPGWEV 280
356 SFLPPGWEMLRILAPNRPFFIDHNKRTTWEDPRILKFPVHMSKSLNPNNDLGPPLPGWEE 415
281 RSTVSGRIYFDNHNRTQFTDPRILHIMNHQCQKESQCLPLPSEGSLEDELPAPORY 340
416 RIHLDGRFTFYIDHNSKITQWEDPRILQ-----PALTG-----PAVPY 452
341 ERDLVQKILVLRHLSLQOPQAGHCRLEVSREEIFEESYRQIMKM-RPKDLKRLMYKFR 399

Db 453 SREPKQYKPYRFXKLKPADINRFEMLKRNINFEESYRIMSVPDVLKARLMEFE 512
 QY 400 GEEGLDYGVAEEMLYLCHENLNPYGLFOYSTDNITWLOINPDSST-NPHHLSYFHFV 458
 Db 513 SEKGIDYGVAAEWEFLLSKEMENPYGJFEYSADNDYTLQINPNSGICNEHLSYFTEFI 572
 QY 459 GIMGLAVHGHYINGGFTVPPYKOLGPKIOLSDLESVDPELHSLWILENDITPVLD 518
 Db 573 GNAGGLAVHGHKLDGFFIRPYKMLGKQITLNDMESVDSYNSLKMILENDITE-LD 631
 QY 519 HFFCEHNAFGRLLOHELKPNGANVPTEENKEVRLVYNNRPMRGLEAQLAQKGFN 578
 Db 632 LNFICIDENFGQYVDLKPNGSEIWTNENKREYIDLVIQRFVNRVQKQNAFLGFT 691
 QY 579 ELIPOLLRPPOKHELIIIGLDKIDLDMKSNTRKHCVADSN-IYVWFQAVETPE 637
 Db 692 ELIPDLIKIPENELIELMCGLDVDVDMROHSYKNGYCPNHEVIMQFKAVALMDA 751
 QY 638 EERARLLQFVGTSTVPLQGFKALOGSTGAAGPRLFTIHLIDANTDNEPKATCFNRIDI 697
 Db 752 EKRIQLQVTGTSRVPNGFALYGSN---GPQLFTIEQW-GSPKLPRAHTECNRLDL 807
 QY 698 PPESEYKYEKLLTVEETCGF 720
 Db 808 PPEYEFEDLREKXLMAVENAGGF 830

RESULT 12

US-08-630-916A-48
 ; Sequence 48, Application US/08630916A
 ; Patent No. 601137
 ; GENERAL INFORMATION:
 ; APPLICANT: Pirozzi, Gregorio
 ; APPLICANT: Kay, Brian K.
 ; APPLICANT: Fowlkes, Dana M.
 ; TITLE OF INVENTION: IDENTIFICATION AND ISOLATION OF NOVEL
 ; TITLE OF INVENTION: POLYPEPTIDES HAVING WW DOMAINS AND METHODS OF USING SAME
 ; NUMBER OF SEQUENCES: 124
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Pennie & Edmonds
 ; STREET: 1155 Avenue of the Americas
 ; CITY: New York
 ; STATE: New York
 ; COUNTRY: United States
 ; ZIP: 10036-2711
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patentin Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/630,916A
 ; FILING DATE: 03-APR-1996
 ; CLASSIFICATION: 435
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: MISROCK, S. LESLIE
 ; REGISTRATION NUMBER: 18,872
 ; REFERENCE/DOCKET NUMBER: 1101-203
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (212) 790-9090
 ; TELEFAX: (212) 896-8864/9741
 ; INFORMATION FOR SEQ ID NO: 48:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 906 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS:
 ; TOPOLOGY: unknown
 ; MOLECULE TYPE: peptide
 ; US-08-630-916A-48

Query Match 31.5%; Score 1225; DB 3; Length 906;
 Best Local Similarity 35.5%; Pred. No. 1,86-112;
 Matches 281; Conservative 115; Mismatches 227; Indels 169; Gaps 19;

QY 28 PFAKIVVDGSGGCHSTDTVKNLTDPKNNQHY-----DLYVGKTDSTITISVNNHKKIKKQG 83
 Db 188 PNGSALTDGS-QLPBSRSSGTAVAPE-NRQOPSTNFGGRS-----RTRHSG 234
 QY 84 AGFLGCVRLISMNISRDKTGQRDLCKINPSDPTDAVQIYVSLQTRD--IGTGGSV 141
 Db 235 AS-----ARTPATGEQSPGASRRRQPKXSGHS 264
 QY 142 VDCRGLLENGETYED-----SGPRPLSCFEE-----PAYTDSGAAAGGN 186
 Db 265 GLANGTVNDEFTATDDEBSVGVCTSPPAAPLSVTNPNTTSLPAPATAEG----- 317
 QY 187 CRFVESPSOD--QRLQARLRNPDVRSLSQT PONRPHGOSPELPE----- 230
 Db 318 EEPSTSGTQQLPA-----AAQAPDALPAGMQRELPNGRVVYVDNHTKTTT 363
 QY 231 -----GYEORTVQGVYFLHTQGVSTWHDPRIRPDIN----- 264
 Db 364 WERPLPQWEMKRTDPRGRFYVDNHTRTTQRTAEYVRVYEQWQSQRNQLOGAMQHS 423
 QY 265 -----SVNDELGLPLPPGWEBSVSGRYFVDNHNRTTQFTDPR----- 304
 Db 424 QRLYQFMSASTDHDPLGPLPQWEMKQD--NGRVYVNNHTRTTQWEDPRTQGMWICPAL 482
 QY 305 -----LHIMNHQCCQLKEBSQPLPLBESGLEDEBELPAQYERDLYQKLYLR 352
 Db 483 PPGWEMKYTESEGVYFVDNHTRTTTFDPRPGFESGTQKQSPGAYDRSFRKXHQFFELC 542
 QY 353 HELSIQCPQACHCEVSEIEFEESYKQIKMRPKDLKRLMWKFRGEEGLDYGVAE 412
 Db 543 HSNAL-----PSHVKISVRQTLFEDSFQIANKMYPDLRLRLYIMGEBGLDVGIARE 598
 QY 413 WLYLCHENLNPYGLFOYSTDNITWLOINPDSINPDHLSYFHPVGRINGLAVHGHYI 472
 Db 599 WFFLSHEVLNPMVCLFEYAGKNYVCLQINPASSINPDHLYFFIGRIFIMALYHGHFI 658
 QY 473 NGGFVPPYKOLLGPKIOLSDLESVDPELHSLWILENDITPV-LDHTFCVBNAPGRI 531
 Db 659 DTGFTLPPYKMLNRPPLTKDLESIDPEFNSIWIKENNEBEGGLELYFQDWEILKV 718
 QY 532 LQHELKPNGANVPTEENKEVRLVYNNRPMRGLEAQLAQKGFNELIPQHLKPPDQ 591
 Db 719 TTHELKEGEGSIRVTEENKEEYIMLTDMRFTRGVEEQTKAFLQFNEVABLEWLRYDE 778
 QY 592 XELELITGLDKIDLDMKSNTRKHCVADSNIYVWFQAVETDEERARLLQFVGTG 651
 Db 779 KELEIMLCMOEIDMSDQKSTIYRHYTKSKQIQWQVQVKEMDNEKRIRLQFVGTG 838
 QY 652 RYVLOGFVALOGSTGAAGPRLFTIHLIDANTDNEPKATCFNRIDIPPESEYKYEKLL 711
 Db 839 RLPVGGFPAELIGSN---GPQKFCIDKVGKET-WLPRSTCTNRLDPLPKSYEQRLRELL 894
 QY 712 TAVEETCGFAVE 723
 Db 895 YAIETEGEGQE 906

RESULT 13

US-08-630-916A-46
 ; Sequence 46, Application US/08630916A
 ; Patent No. 601137
 ; GENERAL INFORMATION:
 ; APPLICANT: Pirozzi, Gregorio
 ; APPLICANT: Kay, Brian K.
 ; APPLICANT: Fowlkes, Dana M.
 ; TITLE OF INVENTION: IDENTIFICATION AND ISOLATION OF NOVEL
 ; TITLE OF INVENTION: POLYPEPTIDES HAVING WW DOMAINS AND METHODS OF USING SAME
 ; NUMBER OF SEQUENCES: 124
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Pennie & Edmonds
 ; STREET: 1155 Avenue of the Americas
 ; CITY: New York

QY 483 OLLGKPIQLSDLESVDPELHKSIVWILE--NDITPVLDHTFCV--BHNAFGRILQHCLKPN 539
 DB 639 KLMGKGLFVLDJSDSPVLYOSIKDLLEYGVNVEDDMITFOISQTNLFGNPMYMDLKEN 698
 QY 540 GRNVPTENKKEYYRLVYNNRFRMGIEAFLALOKGF---NELIPQHLKPFDOKELE 595
 DB 699 GKPIPTNENKKEFVNLSDYILINKSVKQFKAFRGFHMVNTNESPDKYLFPRP---EEIE 755
 QY 596 LIIGGLDKIDNDKMSNTRLK--HCVADSNIVRMQAVETPDEERRARLLQFTGSTTRVP 654
 DB 756 LLIIGSRMLDPQALETTEYDGYTRDSVLIRFEFVIHVSFTDEQKRLFLQFTGTDRAP 815
 QY 655 LOGFKALOGSTGAAGRLFTIHLIDANTDNLPRKATCFNRIDIPRYESYKLYEKLLTAV 714
 DB 816 VGLGLKTKMILAKNGP-----DTRBLPFSHTCFNVLLPPEYSKEXKLERLKAI 865
 QY 715 EETCGFAV 722
 DB 866 TYAKGFGM 873

RESULT 15

US-08-767-942A-21
 ; Sequence 21, Application US/08767942A
 ; Patent No. 6068982

GENERAL INFORMATION:

APPLICANT: Rolfe, Mark
 APPLICANT: Chiu, M. Isabel
 APPLICANT: Berlin, Vivian
 APPLICANT: Darnaghez, Veronique
 APPLICANT: Draetta, Giulio
 APPLICANT: Guillaume, Cottarel
 TITLE OF INVENTION: UBIQUITIN CONJUGATING ENZYMES
 NUMBER OF SEQUENCES: 45
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: FOLEY, HOAG & ELLIOT LLP
 STREET: One Post Office Square
 CITY: Boston
 STATE: MA
 COUNTRY: USA
 ZIP: 02109-2170
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent in Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/767,942A
 FILING DATE: 17-DEC-1996
 ATTORNEY/AGENT INFORMATION:
 NAME: Vincent, Matthew P.
 REGISTRATION NUMBER: 36,709
 REFERENCE/DOCKET NUMBER: MIV-029.04
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 617-832-7000
 TELEFAX: 617-832-1000
 INFORMATION FOR SRO ID NO: 21:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 874 amino acids
 TYPE: amino acid
 TOPOLOGY: 1linear
 MOLECULE TYPE: protein
 US-08-767-942A-21

Query Match 13.7%; Score 530.5; DB 3; Length 874;
 Best Local Similarity 26.1%; Fred. No. 2.7e-43;
 Matches 190; Conservative 117; Mismatches 262; Indels 159; Gaps 26;

QY 98 SRLKDTGYRLDLCKLNPSDT---DAVRGQIVVSLQTRDRIGTG-----GSVVDGR 145
 DB 202 SRIDSSGSGDNMLKLTGSDVSVVDIAIR--RVYTRILSNKIEETAFALNALVYLSPNVECD 260
 QY 146 GLLENEGTIVY-----EDSGRPLSCMEBPAPYTDSTGAAG 183

DB 261 LTYHN---VYSRDPNYINLFIIGMENENLHSPYLEMALPLCKAMSKLP-----LAAQ 311
 QY 184 GGNCRFVESPQDQ-----RLQARLRNPD----- 208
 DB 312 GKILRLMSKYNADIRMMETFOOLITYKYIISNEFNSRLVNEHNSNLVNDDAIYAA 371
 QY 209 -----VRGSLQTPQNRPHQOS--PELPE-----GYEQTTVGQVYFLHTQT 249
 DB 372 KCLMWVYANAVGGEVDTNHNEDEDEPIPESSSELTQELLGERRNKKGLRVDPLETEL 431
 QY 250 GYST--WEDPRIPRD--LNSVNCDELPLRPGWEVRSVTSRGRIYFVHNHRTTOFTPL 305
 DB 432 GVKTLDCRKLPIPEEFINB-----PLNEVLMDKDY--FFKVTENKFSMTCPFI 482
 QY 306 HHIMNHQCOLKEPQPLPLPEEGSLDEDELPAQRYERDLVQKLVRLHLSLOQPOAGHC 365
 DB 483 LNAVTKNGLYYDNR--IRMYSE-----RRITVL--YSLVGGQLNPLY 522
 QY 366 RIEVSRERIEFEES---YKQYKMRPKDKRLMYKFRGEBGLDYGVAREWLYLCHENL 422
 DB 523 RLKYVRDHIIDDAVRLLEMIMENPADLKKQLYVEFBEGQVDGVSKEFQOLVEBIF 582
 QY 423 NPYVGLFQYSTDNITYMLQINPDSSINPDHLSYFHFVGSINGLAVFHGYINGCFVPEFYK 482
 DB 583 NPDIGMFTYD--ESTKLFMFNPSSF---ETEGQFTLIGIVGLAIYNNCILDVHFPVYVR 638
 QY 483 OLLGKPIQLSDLESVDPELHKSIVWILE--NDITPVLDHTFCV--BHNAFGRILQHCLKPN 539
 DB 639 KLMGKGLFVLDJSDSPVLYOSIKDLLEYGVNVEDDMITFOISQTNLFGNPMYMDLKEN 698
 QY 540 GRNVPTENKKEYYRLVYNNRFRMGIEAFLALOKGF---NELIPQHLKPFDOKELE 595
 DB 699 GKPIPTNENKKEFVNLSDYILINKSVKQFKAFRGFHMVNTNESPDKYLFPRP---EEIE 755
 QY 596 LIIGGLDKIDNDKMSNTRLK--HCVADSNIVRMQAVETPDEERRARLLQFTGSTTRVP 654
 DB 756 LLIIGSRMLDPQALETTEYDGYTRDSVLIRFEFVIHVSFTDEQKRLFLQFTGTDRAP 815
 QY 655 LOGFKALOGSTGAAGRLFTIHLIDANTDNLPRKATCFNRIDIPRYESYKLYEKLLTAV 714
 DB 816 VGLGLKTKMILAKNGP-----DTRBLPFSHTCFNVLLPPEYSKEXKLERLKAI 865
 QY 715 EETCGFAV 722
 DB 866 TYAKGFGM 873

Search completed: September 21, 2004, 07:49:17
 Job time : 28.106 secs

This Page Blank (uspio)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 21, 2004, 07:28:49 ; Search time 92.5466 Seconds
(without alignments)
2285.667 Million cell updates/sec

Title: US-10-009-945-4
Perfect score: 4038
Sequence: 1 MSNPGRRNGPKRLTVLC.....EKLYEKLLTAHETGFAVE 748

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :
1: Geneseq_29Jan04: *
2: geneseqp1980s: *
3: geneseqp1990s: *
4: geneseqp2000s: *
5: geneseqp2001s: *
6: geneseqp2002s: *
7: geneseqp2003as: *
8: geneseqp2003bs: *
9: geneseqp2004s: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	4038	100.0	748	AAB31477	Aab31477 Amino aci
2	4031	99.8	748	AAW13384	Aaw13384 Human pro
3	4027	99.7	804	AAU19610	Aau19610 Human dla
4	3001	74.3	722	AAE33724	Aae33724 E3 ubiqui
5	3001	74.3	722	AAB31476	Aab31476 Amino aci
6	2992	74.1	722	AAW79861	Aaw79861 Human pro
7	2775	68.7	514	AAU87301	Aau87301 Novel cen
8	2173	53.8	1035	ABW61120	Abw61120 Drosophi
9	1809	44.8	335	ADC64238	Adc64238 Human SMT
10	1545	38.3	766	AAW13385	Aaw13385 Human pro
11	1530.5	37.9	832	ABP73459	Abp73459 Candida a
12	1453.5	36.0	869	ABJ26104	Abj26104 Aspergill
13	1387	34.3	975	AAW93167	Aaw93167 Human ZGG
14	1370.5	33.9	911	ADB75455	Adb75455 Prostate
15	1365.5	33.6	724	AAW37597	Aaw37597 Novel hum
16	1355.5	33.6	724	ADB49320	Adb49320 Novel hum
17	1354.5	33.5	854	ADB75453	Adb75453 Prostate
18	1354.5	33.5	927	AAV25170	Aav25170 Human KIA
19	1354.5	33.5	927	AAE33725	Aae33725 Human NED
20	1353.5	33.5	995	AAE33719	Aae33719 KIAA0439
21	1353.5	33.5	995	ADB75451	Adb75451 Prostate
22	1353.5	33.5	995	AAE60011	Aae60011 Human pro
23	1353.5	33.5	995	ADD45762	Add45762 Human pro
24	1353	33.5	923	ABG01080	Abg01080 Novel hum
25	1348	33.4	834	AAU77715	Aau77715 Diseased

26	1340.5	33.2	759	6	ABJ25504	Abj25504 Aspergill
27	1338.5	33.1	854	6	ABP58332	Abp58332 Human cel
28	1336	33.1	854	2	AAV30949	Aav30949 Murine E3
29	1333	33.0	949	4	ABW59631	Abw59631 Drosophi
30	1333	33.0	949	4	ABW59856	Abw59856 Drosophi
31	1329	32.9	852	2	AAV30948	Aav30948 Human E3
32	1313.5	32.5	903	6	ABW64210	Abw64210 Angiogene
33	1301	32.2	739	4	ABG16477	Abg16477 Novel hum
34	1301	32.2	739	4	AAW68173	Aaw68173 Atrophin-
35	1301	32.2	739	6	AAE33720	Aae33720 Atrophin-
36	1301	32.2	739	6	ABR41097	Abra41097 Human atr
37	1301	32.2	739	7	ADB98717	Adb98717 Human atr
38	1301	32.2	739	7	ADE82513	Ade82513 Human pro
39	1301	32.2	752	8	ADE77070	Ade77070 Human pro
40	1298.5	32.2	870	4	AAE05495	Aae05495 Human ubi
41	1298.5	32.2	870	6	AAE33722	Aae33722 Nedd-4-1
42	1298.5	32.2	906	7	ADB49242	Adb49242 Novel hum
43	1297	32.1	922	5	ABW50049	Abw50049 Human hom
44	1297	32.1	922	5	ABW50708	Abw50708 Human sig
45	1285.5	31.8	906	2	AAW36795	Aaw36795 Novel hum

ALIGNMENTS

RESULT 1
AAB31477
ID AAB31477 standard; protein; 748 AA.
XX
AC AAB31477;
XX
DT 20-APR-2001 (first entry)
XX
DE Amino acid sequence of a human Smurf2 polypeptide.
XX
KW Smurf1; Smurf2; Smad signal transduction; bone morphogenic protein; BMP;
KW transforming growth factor-beta; human; TGF-beta; chondrogenesis;
KW osteogenesis; blood differentiation; cartilage formation; hair growth;
KW neural tube patterning; retinal development; heart induction;
KW morphogenesis; tooth formation; gamete formation.
XX
OS Homo sapiens.
XX
PN WO200077168-A2.
XX
PD 21-DEC-2000.
XX
PF 12-JUN-2000; 2000WO-US016250.
XX
PR 11-JUN-1999; 99US-0138969P.
XX
PA (UNIV) UNIV NEW YORK STATE RES FOUND.
XX
PA (HSCR-) HSC RES & DEV LP.
XX
PI Thomsen GH, Wrana J;
XX
DR MPI: 2001-071267/08.
XX
DR N-PSDB; AAF24853.
XX
PT Novel isolated Smurf protein useful for inhibiting bone morphogenic
PT protein or tumor growth factor-beta activation pathway, for treating
PT cancer and to block osteogenesis, hair growth, tooth formation.
XX
PS Claim 10; Fig 12; 107pp; English.
XX
XX The present sequence represents a human Smurf2 polypeptide. The
XX specificatation also describes a Smurf1 polypeptide. Smurf polypeptides are
XX negative regulators of Smad signal transduction, and antagonists of bone
XX morphogenic protein (BMP) or transforming growth factor-beta (TGF-beta)
XX signalling pathway. Expression of Smurf1 in a cell is useful for
XX inhibiting a BMP or TGF-beta activation pathway in a cell. Smurf
XX polypeptides are useful for blocking chondrogenesis, osteogenesis, blood
XX differentiation, cartilage formation, neural tube patterning, retinal

Db	361	ECITLVRRYKRDVLQVTKILIKRQLSQQQVAGHCRIEVSREBI	FBESYRQVWKRPKDLWK	420
QY	421	RLMIKRRGEEGLDYGVAFAEWYLILSHENLNPYIGLFOYSDDIYTLQINPDSVNPBH		480
Db	421	RLMIKRRGEEGLDYGVAFAEWYLLSHENLNPYGLFOYSRDDIYTLQINPDSVNPBH		480
QY	481	SYHFHFGRLMGMAVFPGHYITDGFILTFYKQLGKSLTLDMEVLDPDLHNSLWILEND		540
Db	481	SYHFHFGRLMGMAVFPGHYITDGFILTFYKQLGKSLTLDMEVLDPDLHNSLWILEND		540
QY	541	ITGLVDHITFCVEHNAYGEIILQHHLKPNKGSIPVNEENKKEYRYLYNMRFLRGIEAOFLLA		6000
Db	541	ITGLVDHITFCVEHNAYGEIILQHHLKPNKGSIPVNEENKKEYRYLYNMRFLRGIEAOFLLA		6000
QY	601	LQGFNEEVLIPQHLTKPFDEKELELIIQGLKIDVNDKVNTRLKRCPPDSNIYKMFKKAV		6600
Db	601	LQGFNEEVLIPQHLTKPFDEKELELIIQGLKIDVNDKVNTRLKRCPPDSNIYKMFKKAV		6600
QY	661	EFPEDEERRALLQFVTVSSRVPLQGFKALQGAAGRLFTIHQIDACTNNLPKATCFENRI		7200
Db	661	EFPEDEERRALLQFVTVSSRVPLQGFKALQGAAGRLFTIHQIDACTNNLPKATCFENRI		7200
QY	721	DIPPYESYEKLYEKLITALEETCGFAVE	748	
Db	721	DIPPYESYEKLYEKLITALEETCGFAVE	748	
RESULT 3				
AAU19610				
ID	AAU19610 standard, protein; 804 AA.			
XX	AAU19610;			
AC				
DT	04-DEC-2001 (first entry)			
XX	Human diagnostic and therapeutic polypeptide (DTHP) #196.			
KW	Human; receptor; diagnostic; therapeutic; gene therapy; vaccine;			
KM	cell proliferative disorder; Crohn's disease; lymphoma; leukaemia;			
XX	acquired immune deficiency syndrome; AIDS; autoimmune disorder;			
XX	respiratory disorder.			
OS	Homo sapiens.			
XX	WO200162927-A2.			
PN				
XX	30-AUG-2001.			
PD				
XX	21-FEB-2001; 2001WQ-US006059.			
PR	24-FEB-2000; 2000US-0184653P.			
PR	24-FEB-2000; 2000US-0184657P.			
PR	24-FEB-2000; 2000US-0184658P.			
PR	24-FEB-2000; 2000US-0184768P.			
PR	24-FEB-2000; 2000US-0184769P.			
PR	24-FEB-2000; 2000US-0184770P.			
PR	24-FEB-2000; 2000US-0184771P.			
PR	24-FEB-2000; 2000US-0184772P.			
PR	24-FEB-2000; 2000US-0184773P.			
PR	24-FEB-2000; 2000US-0184774P.			
PR	24-FEB-2000; 2000US-0184776P.			
PR	24-FEB-2000; 2000US-0184777P.			
PR	24-FEB-2000; 2000US-0184757P.			
PR	24-FEB-2000; 2000US-0184813P.			
PR	24-FEB-2000; 2000US-0184837P.			
PR	24-FEB-2000; 2000US-0184841P.			
PR	24-FEB-2000; 2000US-0185213P.			
PR	24-FEB-2000; 2000US-0185216P.			
PR	12-MAY-2000; 2000US-0203785P.			
PR	15-MAY-2000; 2000US-0204226P.			
PR	16-MAY-2000; 2000US-0204525P.			
PR	16-MAY-2000; 2000US-0204821P.			
PR	16-MAY-2000; 2000US-0204908P.			

PR	16-MAY-2000;	2000US-0205232P.	
PR	17-MAY-2000;	2000US-0204815P.	
PR	17-MAY-2000;	2000US-0205221P.	
PR	17-MAY-2000;	2000US-0205281P.	
PR	17-MAY-2000;	2000US-0205285P.	
PR	17-MAY-2000;	2000US-0205286P.	
PR	17-MAY-2000;	2000US-0205287P.	
PR	17-MAY-2000;	2000US-0205323P.	
PR	17-MAY-2000;	2000US-0205324P.	
XX			
PA	(INCY-)	INCYTE GENOMICS INC.	
PI	Panzer SR, Spito PA, Banville SC, Shah P, Chalup MS, Chang SC,		
PI	Chen A, D'sa SA, Amesley S, Dahl CR, Dam TC, Daniels SB, Dufour GE,		
PI	Flores V, Fong WT, Greenwalt LB, Hillman JL, Jones AL, Liu TF,		
PI	Roseberry AM, Rosen BH, Russo PD, Stockreiter TK, Daiffo A;		
PI	Wright RJ, Yap PE, Yu JY, Bradley DL, Bratcher SR, Chen W;		
PI	Cohen HU, Hodgson DM, Lincoln SE, Jackson S;		
XX			
DR	WP1: 2001-502867/55.		
DR	N-PSDB; AAS31181.		
XX			
PT	Polynucleotides encoding diagnostic and therapeutic proteins, e.g.		
PT	enzymes, hormones and receptors, useful in diagnostics and therapeutics.		
XX			
PS	Claim 27, Page 512-514; 522pp; English.		
XX			
CC	The invention relates to polynucleotides (I) encoding diagnostic and		
CC	therapeutic (DITHP) polypeptides (II), which include e.g. enzymes, and		
CC	proteins involved in growth and development and receptors. (I) and (II)		
CC	may be used in the prevention, diagnosis and treatment of diseases		
CC	associated with inappropriate DITHP expression. For example, (I) and (II)		
CC	may be used to treat disorders associated with decreased polypeptide		
CC	expression by rectifying mutations or deletions in a patient's genome,		
CC	that affect the activity of the DITHPs, by expressing inactive proteins		
CC	or supplementing the patient's own production of them. (I) and (II) may		
CC	be used to treat diseases, for example, cell proliferative disorder,		
CC	Crohn's disease, acquired immune deficiency syndrome (AIDS), lymphoma,		
CC	leukaemia, autoimmune disorders, and respiratory disorders. Additionally,		
CC	(I) may be used to produce the DITHPs, by inserting the nucleic acids		
CC	into a host cell and culturing the cell to express the protein. (I) and		
CC	its complementary sequences may also be used as DNA probes in diagnostic		
CC	assays to detect and quantitate the presence of similar nucleic acids in		
CC	samples, and therefore which patients may be in need of restorative		
CC	therapy. (II) may also be used as antigens in the production of		
CC	antibodies against DITHPs and in assays to identify modulators of DITHP		
CC	expression and activity. The anti-DITHP antibodies and antagonists may		
CC	also be used to down regulate expression and activity. The anti-DITHP		
CC	antibodies may also be used as diagnostic agents for detecting the		
CC	presence of DITHPs in samples (e.g. by enzyme linked immunosorbant assay		
CC	(ELISA)). AAU9415-AAU9625 represent human diagnostic and therapeutic		
CC	(DITHP) polypeptides of the invention		
XX			
SQ	Sequence 804 AA;		
QY	Query Match	99.7%; Score 4027; DB 4; Length 804;	
Db	Best Local Similarity	99.7%; Pred. No. 0;	
QY	Matches 746; Conservative	0; Mismatches 2; Indels 0; Gaps 0	
1	MSNPGRRNRPVXKRLITVLCANLVKQKFFRLPDPFAKVVYVDSGQCHSTDPVKNLTPK	60	
57	MSNPGRRNRPVXKRLITVLCANLVKQKFFRLPDPFAKVVYVDSGQCHSTDPVKNLTPK	116	
61	KNQHDVLVYKGSQDSTIVSWNHKKIHKKQAGFLCCVLLSNAINRLKDTGYRLDLC	120	
117	KNQHDVLVYKGSQDSTIVSWNHKKIHKKQAGFLCCVLLSNAINRLKDTGYRLDLC	176	
121	GENNDTVARGQIVSLQSRDRIIGTGQGVYDCSRLFDNDLPDQWERRRTASGIIQYLNHT	180	
177	GENNDTVARGQIVSLQSRDRIIGTGQGVYDCSRLFDNDLPDQWERRRTASGIIQYLNHT	236	
181	RTTQWRPRTAPASVSSPGRPLSCVDEVTPLSGNGATCGGSSDPLAERVRSQARRN	240	

```

Db      237 RTQWERPTRRASEYSSPGRPLSCFVDNTEPISGNNMGTCCGSSDPRLAERVRVSQRHN 296
Qy      241 YKSRTHLHTPPDLPEGYQRTTQGOVYFLHTQGVSTWMDPRVPRDLNSINCEELPLP 300
Db      297 YMSRTHLHTPPDLPEGYEQRRTTQGOVYFLHTQGVSTWMDPRVPRDLNSINCEELPLP 356
Qy      301 PGMEIRNTATGVRVVDHNNRTTQFTDPRLSANLHLVLRONOLKDQOQOQVSLCDDT 360
Db      357 PGMEIRNTATGVRVVDHNNRTTQFTDPRLSANLHLVLRONOLKDQOQOQVSLCDDT 416
Qy      361 ECLIVPRYKDLVQKLKTLROELSQOQPOACHCRILEVREBEIFEESSYQVWKKRPKDLMK 420
Db      417 ECLIVPRYKDLVQKLKTLROELSQOQPOACHCRILEVREBEIFEESSYQVWKKRPKDLMK 476
Qy      421 RLMTFRGEEGLDGGVAREWLYLISHMLNPPYGLTQYSSDDIYTTQINPDSVYNEHL 480
Db      477 RLMTFRGEEGLDGGVAREWLYLISHMLNPPYGLTQYSSDDIYTTQINPDSVYNEHL 536
Qy      481 SYFHFVGRIMGVAVFHGHYIDGGFTLPFYKQLLCKSITLDDMELVDEDLHNSLWILEND 540
Db      537 SYFHFVGRIMGVAVFHGHYIDGGFTLPFYKQLLCKSITLDDMELVDEDLHNSLWILEND 596
Qy      541 ITGVLDHTEFCEHNAVGGIILQHELKPNKKSIPVNEENKKEVYRLVYVWRELRGIEAOFLA 600
Db      597 ITGVLDHTEFCEHNAVGGIILQHELKPNKKSIPVNEENKKEVYRLVYVWRELRGIEAOFLA 656
Qy      601 LQKGENEYIPQHLKTFDEKELLELIIQGLKIDVNDKVNTRLGCTPDNSIYKMFYKAV 660
Db      657 LQKGENEYIPQHLKTFDEKELLELIIQGLKIDVNDKVNTRLGCTPDNSIYKMFYKAV 716
Qy      661 EFPDEERARLLQYTGSSRVPLQGFALQGAAPRLFTTHQIDACTNNLPKHAHTCENRI 720
Db      717 EFPDEERARLLQYTGSSRVPLQGFALQGAAPRLFTTHQIDACTNNLPKHAHTCENRI 776
Qy      721 DIPPESEYKLYEKLTLTAIETCGFAVE 748
Db      777 DIPPESEYKLYEKLTLTAIETCGFAVE 804

RESULT 4
AAE32724
ID      AAE32724 standard; protein; 722 AA.
XX
AC      AAE32724;
XX
DT      24-MAR-2003 (first entry)
XX
DE      E3 ubiquitin ligase SMURF1 protein.
XX
KW      Viral infection; lymphosarcoma; human immunodeficiency virus; hepatitis;
KW      poliomyelitis; HIV; measles; protein therapy; E3 ubiquitin ligase;
KW      enzyme.
XX
OS      unidentified.
XX
MO      020290549-A2.
XX
PN      14-NOV-2002.
XX
PD      12-MAR-2002; 2002W0-IB002106.
XX
PR      12-MAR-2001; 2001US-0275224P.
PR      31-JUL-2001; 2001US-0308958P.
PR      07-DEC-2001; 2001US-0340170P.
XX
PA      (PROT-) PROTEOLOGICS LTD.
XX
PI      Greener T, Moskowltz H, Reiss Y, Alroy I;
XX      WPI; 2003-111976/10.
XX      N-PSDB; AAD50458.
XX
PT      New protein complex comprising HECT-RC1, viral maturation scaffolding

```

protein (WMS), and/or HIV gag protein, useful for treating viral infections, such as lymphosarcoma, HIV, hepatitis, poliomyelitis, measles, or Ebola.

Disclosure, Fig 14; 150pp; English.

The invention relates to a method for modulation of viral maturation. The invention also provides an isolated protein complex comprising a HECT-RC1 selected from HECT-1W, HECT-RC1, Gag protein, Gag late domain, P13, actin, myosin, Hsp60, Hsp90, STAM1, STAM2A, STAM2B, VHS-VIM, GTPase, E2 enzyme, teg101, cullin, HRC1, HRC2, HRC3, Nedd4-like protein or clathrin. The complexes, proteins, antibodies and methods are useful for treating viral infections, such as lymphosarcoma, human immunodeficiency virus (HIV), hepatitis, poliomyelitis, measles, or Ebola and for inhibiting budding in a subject. They are also useful in diagnostic assays for determining whether a cell is infected with a virus and for characterizing the nature, progression and/or infectivity of the infection. The invention is also useful in protein therapy. The present sequence is E3 ubiquitin ligase SMURF1 protein used to illustrate the method of the invention.

Sequence 722 AA:

Query Match 74.3%; Score 3001; DB 6; Length 722;
Best Local Similarity 74.6%; Pred. No. 8e-273;
Matches 565; Conservative 63; Mismatches 71; Indels 58; Gaps 8;

```

Qy      12 VETRLTVCAKLVKKDPEFLPDPRAKYVNGSGCHSTDPVXNTLPKXNOHDYITGK 71
Db      4 IKRLTLVCAKLVKKDPEFLPDPRAKYVNGSGCHSTDPVXNTLPKXNOHDYITGK 63
Qy      72 SDSVTISVWNEKKIHKKGAGFLGCVRLLSNAINLKDQYQRLDCLGPNDDTVRGQ 131
Db      64 TDSITISVWNEKKIHKKGAGFLGCVRLLSNAINLKDQYQRLDCLGPNDDTVRGQ 123
Qy      132 IVYSIQSRDRIGTGQGVVDCSRPLPNDLPDQWBERRTASGRIQVNLHTTRITQWBRPFR 191
Db      124 IVYSIQSRDRIGTGQGVVDCSRPLPNDLPDQWBERRTASGRIQVNLHTTRITQWBRPFR 154
Qy      192 ASEYSSPGRPLSCFVDNTEPISGNNMGTCCGSSDPRLAERVRVSQRHN 242
Db      155 --EDSGPGRPLSCFVDNTEPISGNNMGTCCGSSDPRLAERVRVSQRHN 212
Qy      243 ---SRTHLHTPPDLPEGYEQRRTTQGOVYFLHTQGVSTWMDPRVPRDLNSINCEELGP 298
Db      213 QTPQNRPHGHOSPELPEGYEQRRTTQGOVYFLHTQGVSTWMDPRVPRDLNSINCEELGP 272
Qy      299 LPPGEIRNTATGVRVVDHNNRTTQFTDPRLSANLHLVLRONOLKDQOQOQV---S 354
Db      273 LPPGEIRNTATGVRVVDHNNRTTQFTDPRLSANLHLVLRONOLKDQOQOQV---S 328
Qy      355 LCPDDECTLVPRYKDLVQKLKTLROELSQOQPOACHCRILEVREBEIFEESSYQVWKKR 414
Db      329 L---EDELPAQRYERDLVQKLKTLROELSQOQPOACHCRILEVREBEIFEESSYQVWKKR 385
Qy      415 PKDLMKRLMIKFRGEEGLDGGVAREWLYLISHMLNPPYGLTQYSSDDIYTTQINPDSV 474
Db      386 PKDLMKRLMIKFRGEEGLDGGVAREWLYLISHMLNPPYGLTQYSSDDIYTTQINPDSV 445
Qy      475 VNEBELSYFHFVGRIMGVAVFHGHYIDGGFTLPFYKQLLCKSITLDDMELVDEDLHNSL 534
Db      446 INPDLSYFHFVGRIMGVAVFHGHYIDGGFTLPFYKQLLCKSITLDDMELVDEDLHNSL 505
Qy      535 WLENDITGVLDHTEFCEHNAVGGIILQHELKPNKKSIPVNEENKKEVYRLVYVWRELRG 594
Db      506 WLENDITGVLDHTEFCEHNAVGGIILQHELKPNKKSIPVNEENKKEVYRLVYVWRELRG 565
Qy      595 EAOFLAQKGENEYIPQHLKTFDEKELLELIIQGLKIDVNDKVNTRLGCTPDNSIYK 654
Db      566 EAOFLAQKGENEYIPQHLKTFDEKELLELIIQGLKIDVNDKVNTRLGCTPDNSIYK 625
Qy      655 WFAKAVEFDEERARLLQYTGSSRVPLQGFALQGAAPRLFTTHQIDACTNNLP 711
Db      625 WFAKAVEFDEERARLLQYTGSSRVPLQGFALQGAAPRLFTTHQIDACTNNLP 711

```

Db 626 WFWQAVETDEERBARLLQFTVGTSTRVPLQGFKALQSGTGAAPRLFTIHIDANTDMLP 685

QY 712 KAHTCFNRIDIPPYSEYKLYEKLTAIEETGFAVE 748

Db 686 KAHTCFNRIDIPPYSEYKLYEKLTAIEETGFAVE 722

RESULT 5

AA31476

ID AA31476 standard; protein; 723 AA.

XX AA31476;

AC

XX

XX

XX 20-APR-2001 (first entry)

XX

XX Amino acid sequence of a human Smurf1 polypeptide.

XX Smurf1; Smurf2; Smad signal transduction; bone morphogenic protein; BMP;

KW transforming growth factor-beta; human; TGF-beta; chondrogenesis;

KW osteogenesis; blood differentiation; cartilage formation; hair growth;

KW neural tube patterning; retinal development; heart induction;

KW morphogenesis; tooth formation; gamete formation.

XX Homo sapiens.

OS

XX MO200077168-A2.

XX

XX

XX 21-DEC-2000.

XX

XX 12-JUN-2000; 2000MO-US016250.

XX

XX

XX 11-JUN-1999; 9905-0138969P.

XX

XX (UNIV) UNIV NEW YORK STATE RES FOUND.

PA (HSCR-) HSC RES & DEV LP.

XX

XX Thomsen GH, Wraana J;

PI

XX

XX WPI; 2001-071267/08.

DR N-PSDB; AAF24852.

XX

XX

XX Novel isolated Smurf protein useful for inhibiting bone morphogenic

PT protein or tumor growth factor-beta activation pathway, for treating

PT cancer and to block osteogenesis, hair growth, tooth formation.

XX

XX

PS Claim 6; Fig 10; 107pp; English.

XX

XX The present sequence represents a human Smurf1 polypeptide. The

CC specification also describes a Smurf2 polypeptide. Smurf polypeptides are

CC negative regulators of Smad signal transduction, and antagonists of bone

CC morphogenic protein (BMP) or transforming growth factor-beta (TGF-beta)

CC signaling pathway. Expression of Smurf1 in a cell is useful for

CC inhibiting a BMP or TGF-beta activation pathway in a cell. Smurf

CC polypeptides are useful for blocking chondrogenesis, osteogenesis, blood

CC differentiation, cartilage formation, neural tube patterning, retinal

CC development, heart induction and morphogenesis, hair growth, tooth

CC formation, gamete formation and a wide variety of tissue and organ

CC formation processes, and hinder the regeneration, growth, maintenance,

CC etc., of bone and other tissues that are dependent on the BMP pathway.

CC The polypeptide is useful for screening for various drugs and/or

CC antibodies that can either enhance the BMP pathway, or inhibit it

XX

XX

XX Sequence 723 AA;

SO

Query Match 74.3%; Score 3001; DB 4; Length 723;

Best Local Similarity 74.6%; Pred. No. 8e-273;

Matches 565; Conservative 63; Mismatches 71; Indels 58; Gaps 8;

QY 12 VGLRLTVCAKNLYKKDFRLLPDPFAKVVVDGSGQCHSTDTVKNTLDPKMNQHYDLYGK 71

Db 5 IKIRITLVCAKNLAKKDFRLLPDPFAKIVVDGSGQCHSTDTVKNTLDPKMNQHYDLYGK 64

QY 72 SDSVTITSVNHNKHKIKKQAGAGFLGCVRLLSNAINRLKDTGVRDLCLGSDNDMDYRQG 131

Db 65 TDSITISVWNHKKIKKQAGAGFLGCVRLLSNAINRLKDTGVRDLCLGSDNDMDYRQG 124

QY 132 IVVSLQSRDRIGTGGQVVDCSRLLFDNDLPDQWEERRTASRIQYLNHTTTQWESPTPR 191

Db 125 IVVSLQSRDRIGTGGQVVDCSRLLFDNDLPDQWEERRTASRIQYLNHTTTQWESPTPR 155

QY 192 ASEYSSPGRPLSCFVDENTPISGTNGATCG-----QSPDRLAEREVRQRRHNM 242

Db 156 --EDSGGRPLSCFMEBPAEYTDSTGAAGGNCRFESPSQOORLOAQRLKRPDVNGSL 213

QY 243 -----SRTHLPDPLPEGYQORTTOOQGVFLHTQTVSTHMDPRVPRDLSNICEELGP 298

Db 214 QTPQRRPHGOSPLPBGTEORTTVOQGVFLHTQTVSTHMDPRVPRDLSNICEELGP 273

QY 299 LPPGMEIRNTATGRVYFVDHNNRTTFTDPRLSANHLVLRNQOLKQOQOQV----S 354

Db 274 LPPGMEVSTVSGRIYFVDHNNRTTFTDPR-----LHINNHQOLKEPSQPLPSEGS 329

QY 355 LCPDDTECLTVPRKRLVOKLKLROELSGOQOPQAGHCRIVSREEIFESYRQVMKMR 414

Db 330 L--EDELPAQRYERDLVQKLKVRHELSIQDPAQHCRIVSREEIFESYRQVMKMR 386

QY 415 PKDLKRLMIKFRBEGGLDYGVAREWLYLSHEMLNPPYGLFQYSSDIYTIQINDSA 474

Db 387 PKDLKRLMVKFRBEGGLDYGVAREWLYLSHEMLNPPYGLFQYSSDIYTIQINDSA 446

QY 475 VNPEHLSTFHFVGRIMGAVFHGYIDGFTLPFYKQLGKSIITLDMELVDPPLHNSLV 534

Db 447 INPDHLSTFHFVGRIMGAVFHGYIDGFTLPFYKQLGKSIITLDMELVDPPLHNSLV 506

QY 535 WLENDITGVLDHFCVFNHAYGELIIOHELKPNKSIIPVNEENKKEYRLVYNNRPLGI 594

Db 507 WLENDITGVLDHFCVFNHAYGELIIOHELKPNKSIIPVNEENKKEYRLVYNNRPLGI 566

QY 595 EAQFLALQKGFNEVLPQHLKTFDEKELELIICGIGKIDVNDKYNTRKCTDPSNTVK 654

Db 567 EAQFLALQKGFNEVLPQHLKTFDEKELELIICGIGKIDVNDKYNTRKCTDPSNTVK 626

QY 655 WFWQAVETDEERBARLLQFTVGTSTRVPLQGFKALQSGTGAAPRLFTIHIDANTDMLP 711

Db 627 WFWQAVETDEERBARLLQFTVGTSTRVPLQGFKALQSGTGAAPRLFTIHIDANTDMLP 686

QY 712 KAHTCFNRIDIPPYSEYKLYEKLTAIEETGFAVE 748

Db 687 KAHTCFNRIDIPPYSEYKLYEKLTAIEETGFAVE 722

RESULT 6

AA79861

ID AA79861 standard; protein; 722 AA.

XX

XX

XX AA79861;

AC

XX

XX

XX 06-NOV-2001 (first entry)

XX

XX Human protein SEQ ID NO 3507.

DE

XX

XX Human; cytokine; cell proliferation; cell differentiation; gene therapy;

KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;

KW tissue growth factor; immunomodulatory; cancer; leukaemia;

KW nervous system disorder; arthritis; inflammation.

XX

XX

XX Homo sapiens.

OS

XX

XX MO200157190-A2.

PN

XX

XX

XX 09-AUG-2001.

PD

XX

XX

XX 05-FEB-2001; 2001MO-US004098.

XX

XX

XX 03-FEB-2000; 2000US-00496914.

PR 27-APR-2000; 2000US-00560875.

[illegible][illegible]

RESULT 7
A21187201

DT 05-JUN-2002 (first entry)

Novel central nervous system protein #211

KM Centralnervous system; CNS; autoimmune disease; rheumatoid arthritis
KM hyperproliferative disorder; neoplasia; cardiovascular disorder
KM cardiac arrest; cerebrovascular disorder; ischaemia; angiogenesis;
KM nervous system disorder; Alzheimer's disease; AIDS; ocular disorder
KM acquired immunodeficiency virus; dysphagia; gastrointestinal disorder
KM adenocarcinoma; reproductive system disorder; testicular feminization
KM endocrine disorder; diabetes; cancer; leukaemia; neovascularisation;
KM respiratory disorder; renal disorder; kidney failure; blood disorder;
KM myocardial infarction; wound healing; cell proliferation; skin aging;
KM food additive; food preservative; gene therapy.

OS Homo sapiens

PN W0200155318-A2

PD 02-AUG-2001

PF 17-JAN-2001; 2001WO-US001332

PR 31-JAN-2000; 2000US-0179065P

PR 24-FEB-2000; 2000US-0184664P

PR 16-MAR-2000; 2000US-0189874P
ED 17-MAR-2000 010007Z

PR 18-APR-2000; 2000US-0198123P
 DP 19-MAY-2000; 2000US-0205715P

PR 07-JUN-2000; 2000US-0209467P
PR 38-JUN-2000; 2000US-0214885D

PR 30-JUN-2000; 2000US-0215135P
07-JUL-2000; 2000US-0215547P

PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P

PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000IS-0218390P

PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.

14-AUG-2000; 2000US-0224519P.

14-AUG-2000; 2000US-0225214P.

[illegible]

PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249300P.
PR 17-NOV-2000; 2000US-0249305P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 06-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251889P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0253678P.
PA (HUMA-) HUMAN GENOME SCI INC.
XX
XX
FI Rosen CA, Barash SC, Ruben SM;
XX WPI; 2001-581633/65.
DR N-PSDB; ABR43631.
XX
PT New isolated nucleic acid encoding a protein for diagnosing, preventing,
PT treating or ameliorating medical conditions and used as food additives on
PT preservatives.
XX
PS Claim 9; SEQ ID NO 819; 837bp; English.
XX
XX The invention describes an isolated nucleic acid molecule (I) encoding a
XX novel central nervous system protein. (I) and polypeptides (Iii) encoded
XX by (I), are used to treat a medical conditions and in diagnosis of a
XX pathological condition. Disorders which are diagnosed or treated include
XX autoimmune diseases e.g. rheumatoid arthritis, hyperproliferative
XX disorders e.g. neoplasms of the breast or liver, cardiovascular disorders
XX e.g. cardiac arrest, cerebrovascular disorders e.g. cerebral ischemia,
XX angiogenesis, nervous system disorders e.g. Alzheimer's disease and
XX amyotrophic lateral sclerosis, infections caused by bacteria, viruses
XX e.g. Acquired immunodeficiency virus (AIDS) and fungi, ocular disorders
XX e.g. corneal infection, gastrointestinal disorders e.g. dysphagia,
XX adenocarcinomas and irritable bowel syndrome, reproductive system
XX disorders e.g. testicular feminisation, endocrine disorders e.g. diabetes
XX and pituitary dwarfism, cancers and disorders at the cellular level e.g.
XX leukemias, disorders involving neovascularisation e.g. malignancies,
XX respiratory disorders e.g. nonallergic rhinitis, renal disorders e.g.
XX acute kidney failure and blood related disorders e.g. myocardial
XX infarction. The polypeptides can also be used to aid wound healing and
XX epithelial cell proliferation, to prevent skin aging due to sunburn, to

CC maintain organs before transplantation, for supporting cell culture of
CC primary tissues, to regenerate tissues and in chemotaxis. The
CC polypeptides can also be used as a food additive or preservative to
increase or decrease storage capabilities, fat content, lipid, protein,

Query Match 68.7%; Score 2775; DB 4; Length 514;
Best Local Similarity 99.8%; Pred. No. 8.9e-252;
Matches 513; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```

QY 235 SCRRNNYSRTHLHPDLPESGEQRTTQCGVYFLHTQTGVSTWMDPRVRLDININCE 294
DB 1 SRHHNNYSRTHLHPDLPESGEQRTTQCGVYFLHTQTGVSTWMDPRVRLDININCE 60
QY 295 ELGRLPGMEIRNTATGVRVYFDHNNRTQFTDPLSANLHLVNRQQLDQOQOQVVS 354
DB 61 ELGRLPGMEIRNTATGVRVYFDHNNRTQFTDPLSANLHLVNRQQLDQOQOQVVS 120
QY 355 LCPDDTECTVPRYKRDVQKILROELSOQOPAGHCRIEVSREIEFESYRQVWKM 414
DB 121 LCPDDTECTVPRYKRDVQKILROELSOQOPAGHCRIEVSREIEFESYRQVWKM 180
QY 415 PKDLKRLMKRGEGLDYGVARHMLYLSHENTANYGLFOYSRSDITTLQINPDSA 474
DB 181 PKDLKRLMKRGEGLDYGVARHMLYLSHENTANYGLFOYSRSDITTLQINPDSA 240
QY 475 VNPEHLSEYHFGVIRMGAVFEGHYIDGGFTLPFYKQLIGKSTLDEVELVDPILNSLV 534
DB 241 VNPEHLSEYHFGVIRMGAVFEGHYIDGGFTLPFYKQLIGKSTLDEVELVDPILNSLV 300
QY 535 WLENDITGVLDHTCEVHNAYAGELIQHELKPNKSPVNEENKKEVYRLVYVNRFLRGI 594
DB 301 WLENDITGVLDHTCEVHNAYAGELIQHELKPNKSPVNEENKKEVYRLVYVNRFLRGI 360
QY 595 EAQFLALOKGFNEV-PQHLKTFDEKELELLICGKIDVNDVWVNTRLKHCPTDSNIYK 654
DB 361 EAQFLALOKGFNEV-PQHLKTFDEKELELLICGKIDVNDVWVNTRLKHCPTDSNIYK 420
QY 655 WFWKAVEFPDEERRARLLQVYTGSSRVPLQGFKALQGAAGRLFTTHQIDACTNNLPAH 714
DB 421 WFWKAVEFPDEERRARLLQVYTGSSRVPLQGFKALQGAAGRLFTTHQIDACTNNLPAH 480
QY 715 TCENRIDIPYSEYEKLEKLLTALIEETCGFAVE 748
DB 481 TCENRIDIPYSEYEKLEKLLTALIEETCGFAVE 514

RESULT 8
ABB61120
ID ABB61120 standard; protein; 1035 AA.
XX
AC ABB61120;
XX
DT 26-MAR-2002 (first entry)
XX
DE Drosophila melanogaster polypeptide SEQ ID NO 10152.
XX
KW Drosophila; developmental biology; cell signalling; insecticide;
XX pharmaceutical.
XX Drosophila melanogaster.
XX OS
XX WO200171042-A2.
XX
XX 27-SEP-2001.
XX
XX 23-MAR-2001; 2001WO-US009221.
XX
XX 23-MAR-2000; 2000US-0191637P.
XX
XX 11-JUL-2000; 2000US-00614150.
XX
XX (PEKE ) PE CORP NY.
XX
XX Venter JC, Adams M, Li PWD, Myers EW;

```

XX WPI; 2001-656860/75.
DR N-PSDB; ABL05223.
XX
PT New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signalling and cell-cell
PT interactions.

XX Disclosure; SEQ ID NO 10152; 21bp + Sequence Listing; English.

XX The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL01840-ABL016175), expressed DNA
CC sequences (ABL01840-ABL016175) and the encoded proteins (AB057737-
CC AB07072). The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pat_sequences

XX Sequence 1035 AA;

Query Match 53.8%; Score 2173; DB 4; Length 1035;
Best Local Similarity 43.2%; Pred. No. 1.4e-194;
Matches 456; Conservative 104; Mismatches 154; Indels 342; Gaps 16;

```

QY 7 RRRGPKALMLTLCANLVKDFRLPDPFAVYVDGSGQCHSTDTYKNTLDPKKNQHYD 66
DB 8 RRRGPKALMLTLCANLVKDFRLPDPFAVYVDGSGQCHSTDTYKNTLDPKKNQHYD 67
QY 67 LVYKSDSVTLISVWNNHKKIKHKQAGFLGCYVLLSNALNRLYDQYORLDLCKLGPNNND 126
DB 68 LVYKSDSVTLISVWNNHKKIKHKQAGFLGCYVLLSNALNRLYDQYORLDLCKLGPNNND 125
QY 127 TTRGQIVSLQSRD-----IGTGQVVDGSR--LFPNDLPGMEERTASGRIOY 175
DB 126 LVYKSDSVTLISVWNNHKKIKHKQAGFLGCYVLLSNALNRLYDQYORLDLCKLGPNNND 185
QY 176 LNHIRRTQWERTP----- 191
DB 186 VNHAKSTQMDPRQPGVYVSSHATSPOQRANTHNGSGDROAPAGPRTSTCTNLMNNG 245
QY 192 -----ASEYSSPGRPLSCFV----- 206
DB 246 HRSRLSVTASDERHSTELISSVCKENTSPPTPVASATTPGKKTSSNSSSAGRTLEQ 305
QY 207 -----DENTPISGT-----N 216
DB 306 RPTNEPATPTSTTSASVRLHSNDNHVKTPKHQTNGHAPPESTPTSPGQONVYNGNAQN 365
QY 217 GAT-----CGSSDPRLAER----- 231
DB 366 GSTSGNGSGQAAPQPSANGMTQEDATTTSPSTTSPRHSQSPPTNISPPASVTPSA 425
QY 232 -----RVASQHRNRYMGRTHLHP----- 251
DB 426 NGNVHSPNANSTPAGSGGSGSYTATPGORSQRSSQGEESTRRSSRGTRNGTS 485
QY 252 -----DLEGEYQRTTQGOVYFLHTQTGVSTWMDPR 283
DB 486 GGGGGGGGSGQRYSAAIAAQAARPPUDLPFGYEMRTTQGOVYFTHIPGVSTWMDPR 545
QY 284 VPRDL--SNINCEELGRLPGMEIRNTATGVRVYFDHNNRTQFTDPLSANLHLVNRQ 341
DB 546 IPRDFTQHLTLDAIGPLSGMEORKTASGRVYFDHNNRTQFTDPLSANLHLVNRQ 605
QY 342 N-----QLNDQOQOQVVS-----CPDTE 361
DB 606 TVPPTSAANAGTPAPSPATPTPSAAAAPVQATPASNATPTTLTTTNPHRIVPDLIQ 665
QY 362 C-----TVPRYKRDVQKILROELSOQOPAGHCRIEVSREIEFESYRQVWKM 416

```

Db 666 GLEGGADLPKRYRDIYVGLRALRTELQTMQPOSGHCRLEVRNEIPEESYRLIMKRAK 725
 Qy 417 DLAKRMIFKRGEGDYGAVAREWLYLSHEMLNPPYGLFOYSRDDIYTLQINPDASN 476
 Db 726 DLAKRMIFKRGEGDYGAVAREWLYLSHEMLNPPYGLFOYSRDDIYTLQINPDASN 785
 Qy 477 PEHLISYFHEVGRIMGNVAFHGHYIDGGFTLLPYKQLGKSTLLDMELVDPJLANSWTI 536
 Db 786 PDHLISYFHEVGRIMGNVAFHGHYIDGGFTLLPYKQLGKSTLLDMELVDPJLANSWTI 819
 Qy 537 LENDIGVADHPCVEHNAVAGELIIOHELKPNKSPVNEENKKEYRLVLMVRMRPIGTEA 596
 Db 820 LENDIGVADHPCVEHNAVAGELIIOHELKPNKSPVNEENKKEYRLVLMVRMRPIGTEA 879
 Qy 597 QFLALQKGFENEVLPQHLKTFDEKELELLICGLKIDVNDKVNTRLKCTPDNSITVWF 656
 Db 880 QFLALQKGFENEVLPQHLKTFDEKELELLICGLKIDVNDKVNTRLKCTPDNSITVWF 939
 Qy 657 WKAVEFDEERRARLLQFTGSSRVPLQGFKAQ---GAAGRLFTIH-QIDACTNMLPK 712
 Db 940 WKAVEFDEERRARLLQFTGSSRVPLQGFKAQ---GAAGRLFTIH-QIDACTNMLPK 999
 Qy 713 AHTCFNRIDIPPESEYKLYEKLTAIEETGFAVE 748
 Db 1000 AHTCFNRIDIPPESEYKLYEKLTAIEETGFAVE 1035
 RESULT 9
 ADC64238
 ID ADC64238 standard; protein; 335 AA.
 AC ADC64238;
 XX 18-DEC-2003 (first entry)
 DT Human SMURF2_v2.
 DE
 XX Hepatotropic; vaccine; protein-protein interaction;
 KM Transforming Growth Factor beta; TGF beta; hepatitis;
 KW Selected Interacting Domain; SID; Dalt; human; SMURF2_v2.
 XX Homo sapiens.
 OS
 PN WC02003045990-A2.
 PD 05-JUN-2003.
 PF 26-NOV-2002; 2002MO-EP013866.
 XX 26-NOV-2001; 2001US-0333348P.
 PR 31-MAY-2002; 2002US-0384537P.
 PR 30-OCT-2002; 2002US-0422471P.
 PA (HYBR-) HYBRIGENICS.
 XX Legrain P, Gauthier J, Colland F, Jacq X;
 PI WPI; 2003-505185/47.
 DR N-PSDB; ADC64442.
 XX New complex between two interacting proteins, useful for screening
 PT molecules that inhibit transforming growth factor beta (TGF beta) or TGF
 beta super-family of cytokines pathway for diagnosing or treating TGF
 beta diseases or disorders.
 PS Disclosure; SEQ ID NO 22; 148bp; English.
 CC The present invention relates to protein-protein interactions and
 CC complexes involved in Transforming Growth Factor (TGF) beta disorders
 CC and/or diseases. The complex between two interacting proteins is useful
 CC for screening molecules that inhibit TGF beta for diagnosing or treating
 CC diseases or disorders involving TGF beta e.g., hepatitis. To illustrate
 CC the invention, selected interacting domains (SID) of proteins and their

CC coding sequences were isolated (ADC64238-ADC64318). Sequences ADC64318-
 CC ADC64424 are bait sequences used to isolate the SIDs of the invention.
 CC Sequence 335 AA;
 SQ
 Query March 44.8%; Score 1809; DB 7; Length 335;
 Best Local Similarity 99.1%; Pred. No. 4,3e-161;
 Matches 332; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 Qy 1 MSNRRRRNGPVKRLTYLCAKLVKDFRLLPDPFAKVVVDSGSGCHSTDTVKNITDLPK 60
 Db 1 MSNRRRRNGPVKRLTYLCAKLVKDFRLLPDPFAKVVVDSGSGCHSTDTVKNITDLPK 60
 Qy 61 WNOHYDLYIGSDSVTISVWNNKKIKKQAGFLGCVRLSNAINRLKDTGYORLDCKL 120
 Db 61 WNOHYDLYIGSDSVTISVWNNKKIKKQAGFLGCVRLSNAINRLKDTGYORLDCKL 120
 Qy 121 GPNDNDIVRQIVVLSQSRDRTIGTGVVDCSRLLFNDLDPGMEERRTAGRIQYLNHIT 180
 Db 121 GPNDNDIVRQIVVLSQSRDRTIGTGVVDCSRLLFNDLDPGMEERRTAGRIQYLNHIT 180
 Qy 181 RTTQWERPTRPASISYSGFRLSCFVDENPTISGTNGATGQSSDPRLARVRNSQSHRN 240
 Db 181 RTTQWERPTRPASISYSGFRLSCFVDENPTISGTNGATGQSSDPRLARVRNSQSHRN 240
 Qy 241 YMSRTHLHTPPDLPEGYRQRTTQOGQVYFLHTGTGVTMHPRVPRDLNSINCELGPLP 300
 Db 241 YMSRTHLHTPPDLPEGYRQRTTQOGQVYFLHTGTGVTMHPRVPRDLNSINCELGPLP 300
 Qy 301 PGWEIRNTATGRVYFVDANNRTTQFTDPRLSANLH 335
 Db 301 PGWEIRNTATGRVYFVDANNRTTQFTDPRLSANLH 335
 RESULT 10
 AAW13385
 ID AAW13385 standard; protein; 766 AA.
 AC AAW13385;
 XX 10-JUL-1997 (first entry)
 DT Human protein ubiquitin ligase pub2.
 DE
 XX Protein ubiquitin ligase; pub2; cdc25 phosphatase; CDK kinase; p53;
 KM cell cycle; transgenic animal.
 XX Homo sapiens.
 OS
 PN WO9712962-A1.
 PD 10-APR-1997.
 PF 04-OCT-1996; 96WO-US015930.
 PR 04-OCT-1995; 95US-00539205.
 PR (COLD-) COLD SPRING HARBOR LAB.
 PA Beach D, Caligiuri M, Nefsky B;
 PI WPI; 1997-226206/20.
 DR N-PSDB; AAT47041.
 XX Human and Saccharomyces pombe protein ubiquitin ligase(s) - involved in
 PT cdc25 phosphatase and p53 ubiquitination, and regulate cell growth and
 PT proliferation.
 PS Claim 1; Page 80-84; 108bp; English.
 CC Human protein ubiquitin ligases pub1 (AAW13384), pub2 (AAW13385) and pub3
 CC (AAW13386) are homologues of fission yeast pub1 (AAW13387) and were
 CC identified from cDNA clones (AAT47040-42) obt'd. e.g. from a keratinocyte

CC cDNA library. Pub polypeptides can be produced in transfected host cells.
 CC They can control the steady state level of cdc25 phosphatase, the degree
 CC of CDK kinase (e.g. cdc2) phosphorylation and the steady state level of
 CC p53 (controlling the degree of cell cycle regulation of p53). They can
 CC regulate cell or tissue differentiation, or cell growth or proliferation
 CC by affecting other proteins, can be a specific (ant)agonist of wild-type
 CC protein function and may be used as immunogens to elicit a specific
 CC immune residue
 CC XX

SQ Sequence 766 AA;

Query Match 38.3%; Score 1545; DB 2; Length 766;
 Best Local Similarity 40.8%; Pred. No. 1,2e-135;
 Matches 336; Conservative 121; Mismatches 234; Indels 132; Gaps 18;

QY 1 MSNPRRNGVVKRLTLCAKLVKDFRLPPPAKVVDGSGGCHSDTYVNTIDPK 60
 DB 1 MSNSAQSR---RIRVIIVADGLYKDVFRFPDPFAVLTVDGE-QHTTAAIKKTLNIPY 55
 QY 61 WNOHYDLYIGKSDSVTISVMNHKKIKKQAGFGCVRLISMALNRKIDGYGRL--DLG 118
 DB 56 WNETFEVAVTDSNTAIGVFPQKFF-KKKGGFGVNLNRGVLDLALIGDEMLIDLK 114
 QY 119 KLGPVNDVTGQIVYSIQ----- 137
 DB 115 KL--NENTVVGKTIIMLSTIAQLTLGVSSAASAGARTSTINDPQSSKSSVSHPA 172
 QY 138 -----SRDRIGGGGVVDCSRLFDN-----DLPGWERRRATSGRIQVNHITRTQ 184
 DB 173 SSRAGSPTRDPAAPASPSSEPTFFSFEQDYGRLPGWECTDNLGRITVVDNHTSIT 232
 QY 185 WERTP-----RPASEVSSPGRPLSCVDENTPISGNGATCGGSSDPLAERRVRSQR 238
 DB 233 WIRNPLSSVAGAAAEHLSSAS--SAVTEGVCPSSSNA-----RTEASVLT-- 279
 QY 239 RNVSRTHLTHPPPLPGSYEORTQOGQVFLHTQGVSTHMDPRVPRDLSNIN----- 292
 DB 280 ---SNATTAGGELLPEMBQRTPEGRPFYVDNHTVTTTWDPRQOQVIRSYGGPNAT 335
 QY 293 -----CEELGFLPGWEIRNATATGTVFVDNHNRTTOFTPRLSANLHLNRONOLKQ 347
 DB 336 IQQGVQGLFPLPGWEMRLNLTARVYFVDNHTKTTTWDPRPLPSL-----DQ 384
 QY 348 QQQQVSLCPDTECLVPRYKRLVOKLILRQELSQ--QPOAGHCRLEVSREELFEE 405
 DB 385 -----NVPOYKDRFRRLTYF--LSQPALHPLPGQCHIKVARNHIFED 425
 QY 406 SYRCWTKRPKDLMKRLMKIFRGEGGLDYGVAREMLYLISHMLNPLYGLPOYSRDIY 465
 DB 426 SYAEMQASATDLKRLMKIFRGEGGLDYGLSREYFPLSHENKPFYCLFEYSVDNY 485
 QY 466 TLQINPDGAVNPEHLSTFHFVGRIMGAVPHGHYIDGGFTLLPYKQLGKSTLDMEV 525
 DB 486 TLQINPSGGINPEHLNPFKEIGRVIQGLAIFHRFVDVAFVVSFYKMLQCKVTLQDMES 545
 QY 526 DPDLNSLWTLIENDIGVLDHTECVENHAYGEIQLHELPKNSKIPVNEENKEVRLY 585
 DB 546 DAETRLWLDNDIGVLDITSVNDCGFEVITDLKPNGRNIEVTEENKREYVDLY 605
 QY 586 VNWFFLGLIAQFLALQGFENEVIPOHLTPEKELELLICGLIKIDVNMKVNTRLKH 645
 DB 606 TVM-IQRIEIQGFAYAFEGFSELIPOELINVFDERELHLLIGISEIDMEDAKKHQDYS 664
 QY 646 CTSPNSIVKMFVKAVERFEDEERARILQFTGSSRVVLOQFRLQGAAGRLEFTIQIDA 705
 DB 665 YSENDQILIKFWELMDWSNEKSRRLQFTGTTSRLPVNFGKOLQSDGPRKFTIKAGE 724
 QY 706 CTNNLPKATCFNRIDIPYSEYEKLEKLTALIEETGFAVE 748
 DB 725 -PNKLPRKATCFNRIDLPRYTSKXDLHKLSTIAVEETIGRGE 766

RESULT 11

ABP73459
 ID ABP73459 standard; protein, 832 AA.
 AC ABP73459;
 DT 30-JAN-2003 (first entry)
 DE Candida albicans essential protein SEQ ID NO 7296.

KW Fungus; yeast; tetracycline; promoter; GRACE strain; biosynthesis;
 KM signal transduction; DNA replication; cell division; growth;
 XX proliferation; Candida albicans; fungicide; antifungal.

OS Candida albicans.

PN MO200253728-A2.

PD 11-JUL-2002.

PF 26-DEC-2001; 2001MO-US049486.

PR 29-DEC-2000; 2000US-0259128P.

PR 20-FEB-2001; 2001US-00792024.

PR 22-AUG-2001; 2001US-0314050P.

PA (ELIT-) ELITRA PHARM INC.

PI Roemer T, Jiang B, Boone C, Bussey H, Ohlsen KL;

DR WPI; 2002-566694/60.

DR N-PSDB; ABZ32009.

PT Constructing strains for identifying gene products as effective targets
 for therapeutic intervention, by inactivating in the strain one allele of
 a gene and placing other allele of the gene under conditional expression.

XX Claim 44; SEQ ID NO 7296; 167bp + Sequence Listing; English.

XX The invention relates to constructing (M1) a strain of diploid fungal
 CC cells in which both alleles of a gene are modified, comprising modifying
 CC one allele by insertion or replacement by a cassette having an
 CC expressible selectable marker and modifying other allele by
 CC recombination, of a promoter replacement fragment with a heterologous
 CC promoter, so that expression of the second allele is regulated by the
 CC promoter. (M1) is useful for constructing a strain of diploid fungal
 CC cells in which both alleles of a gene are modified. The diploid fungal
 CC cells having both alleles modified are useful for identifying a gene that
 CC is essential to the survival or growth of a fungus, a gene that
 CC contributes to the virulence and/or pathogenicity of a fungus, a gene
 CC that contributes to the resistance of a diploid fungus to an antifungal
 CC agent, an antifungal agent that inhibits the growth of a diploid fungus
 CC and for identifying a therapeutic agent for treatment of a mammalian
 CC disease. (M1) is useful for identifying a compound which modulates the
 CC activity of a gene product, preferably enzymatic activity, carbon
 CC compound catabolism, biosynthesis, transporter, transcriptional,
 CC translational, signal transduction, DNA replication and cell division
 CC activity. The method is useful for identifying a compound having the
 CC ability to inhibit growth or proliferation of C. albicans cells and for
 CC treating infection by C. albicans. The present sequence is that of an
 CC essential Candida albicans protein used in the method of the invention.
 CC Note: The sequence data for this patent is not represented in the printed
 CC specification but is based on sequence information supplied to Derwent by
 CC the European Patent Office

XX Sequence 832 AA;

Query Match 37.9%; Score 1530.5; DB 5; Length 832;
 Best Local Similarity 38.4%; Pred. No. 3.2e-134;
 Matches 338; Conservative 120; Mismatches 228; Indels 195; Gaps 21;

QY 6 RRRGVPVRLTLVCAKLVKDFRLPPPAKVVDGSGGCHSDTYVNTIDPKNQH 65
 DB 9 RSNNTTINVKVVAESLYKRDVFRQDPFAVLTVDG-QYKTTIRAKTLNPNWNETF 67


```

QY 153 -----RLPND---LPGWEERTASGRIOYLNHTRTQ 184
Db 251 PSSTAAPASAGAAGVNSHOSRTNLSFESQGRLLPAGNEREDNIGRIYYVDHNTTRTT 310
QY 185 WERTPAPASYSFGRPLSCFVDENTPIGNTGATGSSDPLAERRVRSQRHNYM-- 242
Db 311 W---TRPSSNYNEHAQ-----RSOREANMQERSAHQSRMLP 344
QY 243 -----SRTHLTPP-----DLEGEQRTTQOQVY 268
Db 345 EDTGTANSRLPSSSQAHATPPAGGSANAVMMATGATTAGTGGELPFGWEQRTTEGGRPY 404
QY 269 FLHTGTGVSTWHDPRVRDL---SNIN-----CEELGPLPFGWEIRNTATGAVY 314
Db 405 FVDHNTTRTTWVDFRQOYIRMYGQVANGNTNTTIOQOVPQSGPLPFGWEIRNTATVAVY 464
QY 315 FVDHNNRTQTFDPRLSANLHLVLRQNLKQOQOQVSLCPDTECLTPYRYADVLQ 374
Db 465 FVDHNTKITTTWDPRLPSSL-----DQ-----GVFOYKRDPFR 497
QY 375 KXKILNQELSQOOP---QAGHCRLEVSREEIFEESYQVMKRPKDLMKRLMIFRGE 430
Db 498 KLIYFR---SQPALRIMSGQCHVKVRNMIFESYAEIMRQASDLKRLMIFRGE 552
QY 431 GLDYGVARBEWLYLSHEMLNRYGLFQYSRDDITTLQINPDSAVNPHLSYFHFVGRIM 490
Db 553 GLDYGGLSGEFPFLSHENENPFYCLFEYSADNTYTLQINPDSGVNPHLNYFKRIGRV 612
QY 491 GNAVFHGHYIDGFTLFFYKOLLGKSTLDPMELVDPDLHNSLYWLENDITGVLDPHFC 550
Db 613 GLAIFHRRLDSFFIGAFYKWLKRVSLQMEGVEDLHNRLLTWTWMDLIEGVELTFS 672
QY 551 VEHNAVGEIQLHELPNGKSIPIVNEENKEVYRLVYNNRPLRGIEAOLFALQKFNVIY 610
Db 673 VDDEKFGERRITDLRGGRDIPVTENKAEYRLVTEWKIVKRVEEQNAFMSGFNLIP 732
QY 611 QHLTKFPEKELELIIICGLKIDVDMKVNRLKCHCTDSIVVFKWAVYEFBEKRR 670
Db 733 ADLVVVFPERBELIIGIADIIVDMKCHTDYGYQESDEVIONFWKIVSWDAEKSR 792
QY 671 LLOPVYTGSSRVPLQGFKAJOGAAGRPLFTIHOIDACTNNLPEKATCFNRIDIPYSEYK 730
Db 793 LLOFVTGTSRIPIVNGFKDLQSGDGRRTIEK-SGDPAALPKSHGTGNRDLPPYKYET 851
QY 731 LYEXLTAIETCGPAVE 748
Db 852 LEHKNSTAVETLGGQE 869

RESULT 13
AAM93167
ID AAM93167 standard; protein, 975 AA.
AC AAM93167;
XX
DT 24-MAY-1999 (first entry)
XX
DE Human ZGGBP1 protein.
XX
KM ZGGBP1; bipolar affective disorder type I; human; neurological disorder;
KW regulate; antisense; treatment; disease; stroke; dementia; renal;
XX hypertension; nephrosis; cardiovascular.
XX Homo sapiens.
OS
FH Key 1. 975 location/Qualifiers
FT Protein /label= ZGGBP1
FT /note= "Partial sequence"
FT Misc-difference 455
FT /note= "unknown"
PN MO906539-AA1.

```

```

XX 11-FEB-1999.
PD 28-JUL-1998; 98WO-GB002259.
XX
PF 01-AUG-1997; 97GB-00016162.
XX
PR (ZENE ) ZENEGA LTD.
PA
XX Flannery AV, Flannegan MCW;
PI WPI; 1999-153769/13.
XX N-PSDB; AAX22554.
DR
XX New isolated gene associated with neurological disorders - used to
PT develop products for treating e.g. stroke, dementia, renal disorders,
FT hypertension or cardiovascular disorders.
XX
PS Claim 1, Page 42-46; 58pp; English.
XX
CC This sequence represents a human ZGGBP1 protein which is associated with
CC affective neurological disorders such as bipolar affective disorder type
CC I. The invention also describes the isolation of the corresponding mouse
CC gene. The ZGGBP1 protein and homologues or fragments may be used to
CC generate substances which selectively bind to it and in so doing regulate
CC the activity of the protein. The antisense DNA or antisense RNA may be of
CC use in the treatment of diseases or disorders in humans in which the over
CC - or under-regulated production of the gene product has been implicated,
CC e.g. neurologic (stroke, dementia), renal (hypertension, nephrosis),
CC cardiovascular disorders
XX
SQ Sequence 975 AA;
Query Match 34.3%; Score 1387; DB 2; Length 975;
Best Local Similarity 34.4%; Pred. No. 1,4e-120;
Matches 330; Conservative 125; Mismatches 247; Indels 258; Gaps 26;
QY 14 LRLTVLCANLVKQDFRLLDPFAVY---VVDGSGCH--SDYVKTLDPRKNOHYDLY 68
Db 42 LRVKVVSGIDLAKKQIFGASDPYVLSLVADENELALVQTKTKILNPNKNEEFYFR 101
QY 69 IKSQD-SVTISVWNEKTKHKQAGFLCCVR-----LLSNAIN 105
Db 102 VNPNSHRLLEFVFDNRLTRD---FLQGVVPLSLHPTEDPTMERPYTFKDFLLRPSH 158
QY 106 RLKDTGYORLDCCKGPNNDPTVAGQIVVLSQSRDRIQTGGQVDCSLFPND----- 158
Db 159 KSRVKGFLRLKAVYPRKG-----QDEENSDQDDMHGWEVVD-----SNDASQHOE 208
QY 159 -----LPGWEERTASGRIOYLNHTRTQWERPT----- 189
Db 209 ELPPPLPPGWEKVDNIGRTYYVNNRRTQWHRPSLMDVSSSDNNIRQINQEAHR 268
QY 190 -----RPASEYSP-----GRPLSCFVDENTPISGTN-----GATGQSSDPR-- 227
Db 269 FRSRHHISEDLPEPSBGVDPEPWEY-ISEEVNIAGSLGVLLPPPPASGSRISPOL 327
QY 228 -----LAER-----RVRS-----QHRRYMYSRTHLHTPPDL 253
Db 328 SEELSRRLQITPDSNGEQFSSLIQRPBSRRRSCSVYDAVAEQHLPPPSVAVYHTTFL 367
QY 254 PEGYQRTQOQGVYFLHTQGVSTWHDP-----RVPRDL 288
Db 388 PSGWEERKDAKGRYYVNNHNRRTTWTFRPIQALADGASGATNSNNHLTEPQIRRRSL 447
QY 289 SNINCEELGP-----LPGWEIRN 307
Db 448 SSPVYTLXAPLEGAKDPPVRAVVDLTLSNPQSPSPSYNPKQHKYQSGFLPPGWEIRN 507
QY 308 TATGRVYFVDHNNRTTFTDPRLSANLHLVLRQNLKQOQOQVSLCPDD----- 359
Db 508 APNGRPFIDHNTKITTWEDPRLKFPVHM-----SKSTSLNENDIGPLPGW 554

```

```

QY 360 -----TECLTVP--YKRDLYOKLILROELSQOOPQ 389
DB 555 EERIHLDGRTFYIDHNSKITQMEDPLQNPALITGPAVPSRFRKQYDFRKKLLKPADI 614
QY 390 AGHCRLEVSREEIFEESYSQVNMKRPDKLMK-RLMIKFRGEEGLDYGVAREWYLLSHE 448
DB 615 PNRFEWKLHRRNNIFEESYRRIMSVKRPDYLKARLWIEFESEKGLDYGVAREWYLLSHE 674
QY 449 MNPYVGLQYSDDIYTLQINPDNAV-NPEHLSYFHFVGRIMGVAVEHGHYIDGGFTLP 507
DB 675 MNPYVGLQYSDDIYTLQINPDNAV-NPEHLSYFHFVGRIMGVAVEHGHYIDGGFTLP 734
QY 508 FYKOLIGKSTILDDMLVDPDLNLSLWILENDITGVLDHTFCVEHNAVGEIIQHELKEN 567
DB 735 FYKMWLGKQITLNDMSVDSEYNSLKLWILENDPT-ELDLMFCDIBENFGQTYQVDLKEN 793
QY 568 GKSIPVNEBKKEVYVLYNNRFLRGLIEQFLALQKGFNEVYPOHLLKTFDEKELELLIC 627
DB 794 GSEIWTNENKREYIDLVIQWRFVNSVQKMAFLBGFTELLPIDIKLFDENBELLMC 853
QY 628 GLGKIDVNDKVTTRLK--HCTPDSNIYKMFKAVEFPDEERRARLLQFVTGSSRYVLQ 685
DB 854 GLGDVAVNDKROHSIKNGYC-PNHVFIQWFKAVLLMDAEKIRLLQFVTGSRVPMNG 912
QY 686 FKALQAGAPRLFTTHQIDACTNNLPKATCFNRIDIPPEYSYEKLYEKLTAIEETGCF 745
DB 913 FAEVYSGNGPQLFTIEQWGS-PEKLPRAHCFNRIDLPEYTFEDIREKILMAVENAOGF 971

RESULT 14
ADB73455
ID ADB74455 standard; protein, 911 AA.
AC ADB75455;
XX 04-DEC-2003 (first entry)
DE Prostate cancer marker protein.
XX Prostate; cancer; cytostatic; gene therapy; marker.
XX Homo sapiens.
XX W02003009814-A2.
XX 06-FEB-2003.
XX 25-JUL-2002; 2002WC-US023913.
XX 25-JUL-2001; 2001US-0307982P.
XX 22-AUG-2001; 2001US-0314356P.
XX 25-SEP-2001; 2001US-0325020P.
XX 12-DEC-2001; 2001US-0341746P.
XX 05-MAR-2002; 2002US-0362158P.
XX (MILL-) MILLENNIUM PHARM INC.
XX Schlegel R, Monahan JF, Endege WO, Gamavarapu M, Gorbacheva B,
XX Heerth S, Kamakar S, Womsey AM, Glatc K, Zhao X, Anderson D,
XX WPI; 2003-248033/24.
XX New nucleic acid molecule, useful for diagnosing or treating prostate
XX cancer.
XX Claim 4, SEQ ID NO 279; 99P; English.
XX The invention relates to newly discovered cancer markers associated with
XX the cancerous state of prostate cells. Also disclosed is a method of
XX assessing whether a patient is afflicted with prostate cancer. The method
XX of the invention involves assessing whether a patient is afflicted with
XX prostate cancer by comparing the level of expression of a marker in a
XX patient sample and the normal level of expression of the marker in a

```

```

CC control non-prostate cancer sample, where a significant increase in the
CC level of expression of the marker in the patient sample and the normal
CC level indicates that the patient is afflicted with prostate cancer.
CC Nucleic acids of the invention are useful for diagnosing or treating
CC prostate cancer, and may be useful in gene therapy. Sequences given in
CC ADB75177-ADB75631 represent marker cDNA and proteins. Note: The sequence
CC data for this patent did not form part of the electronic specification, but
CC was obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 911 AA;
Query Match 33.9%; Score 1370.5; DB 7; Length 911;
Best Local Similarity 36.7%; Pred. No. 4.3e-119;
Matches 324; Conservative 119; Mismatches 234; Indels 207; Gaps 28;
QY 24 LYKSDFFRLP--DPPAKV--VVDGSGQCH--STDVXKTLDPKRNQHYDLIGKSD-S 74
DB 69 LPTSDELFLPGLCDPYKLSLYVADENBELALVOTKIIKTLNPKMNEEFYFRVPSNRH 128
QY 75 VTISVNMKKIHKKGAGFLGCVR-----LISNAINRLKDTGY 112
DB 129 LLEFVFDENRLTRDD--FLGOVDVPLSHLPTEDPTMERPYTFKDFLLRPSHRKRVKGF 185
QY 113 QRLDCKLGPNDNTVAGQIVLSIQSRDRIGTGQGVYDCSRLFND-----L 159
DB 186 LRLKAVYPPKNG--GDENSDQRDMEHGEVVD--SNDASQHQBELPPPL 235
QY 160 PDGWEERRTASGRLOLYNHTTRTOWERPT-----RPA 192
DB 236 PRGMEKXVDNIGRTYVYVNNHNRITQWHRPSLMDVSSSDNNRIQINQEAHRRFRSRHI 295
QY 193 SYSSSP-----GRPLSCVDENTPISGTN-----GATGQSSDPR-LAE--RR 232
DB 296 SEDLEPESGADVPEPWT-ISEEVNIAGDSLGLAPPSPASGSRISPOELSEUSRR 354
QY 233 VR-----SORHNNYSRCTL-----HTPPD----- 252
DB 355 LQITPDSNGEOPSSLIQRPSSRLSCSVTDAVAEQHLPFGAKDSPVRAVKDTLSNPQ 414
QY 253 -----LPEGEQRTQOQGVYFHTQTVSYTHDPRV-----RD 287
DB 415 SPQEPSPVSPKQHKVQSFPLPGMEWELAPNGRPFPLDHTKTTTIEDPRLKPPVHRS 474
QY 268 LSNICEELGRLPQMEWELRNATAGRVYVVDHNNRTQFTDPRLSANLHLVNRQNLQDQ 347
DB 475 KTSLNPNDLGRLPQMEWELRNATAGRVYVVDHNNRTQFTDPRLSANLHLVNRQNLQDQ 519
QY 348 QQQQVSLCPDTECLTVP--YKRDLYOKLILROELSQOOPQAGHCRLEVSREEIFE 405
DB 520 -----PATGPAVPSRFRKQYDFRKKLLKPADIPNRFEMTLHRRNIFEE 566
QY 406 SYRQWKKRPPDKLMK-RLMIKFRGEEGLDYGVAREWYLLSHEMLNYYGLFCYSRDDI 464
DB 567 SYRRIWSVKRPDYLKARLWIEFESEKGLDYGVAREWYLLSHEMLNYYGLFCYSRDDI 626
QY 465 YTLQINPDNAV-NPEHLSYFHFVGRIMGVAVEHGHYIDGGFTLPFYKOLIGKSTILDDME 523
DB 627 YTLQINPDNAV-NPEHLSYFHFVGRIMGVAVEHGHYIDGGFTLPFYKOLIGKSTILDDME 686
QY 524 LVDPDLNLSLWILENDITGVLDHTFCVEHNAVGEIIQHELKPNKKSIPVNEBKKEVYR 583-
DB 687 SVDSSEYNSLKLWILENDPT-ELDLMFCDIBENFGQTYQVDLKPNGSEIWMNENREYID 745
QY 584 LYVNNRFLRGLIEQFLALQKGFNEVYPOHLLKTFDEKELELLICGLKIDVNDKVTTRL 643
DB 746 LVQWRFVNSVQKMAFLBGFTELLPIDIKLFDENBELLMCGLGDVAVNDKROHSIY 805
QY 644 K--HCTPDSNIYKMFKAVEFPDEERRARLLQFVTGSSRYVLQFKALQAGAPRLFTIH 701
DB 806 KNGYC-PNHVFIQWFKAVLLMDAEKIRLLQFVTGSRVPMNGFAEYSGNGPQLFTIE 864
QY 702 QIDACTNNLPKATCFNRIDIPPEYSYEKLYEKLTAIEETGCF 745

```



```

DB      865 QWGS-PEKLPRAHTCFNRLLDPFYETFEDELKELMLAVENAGSF 907
RESULT 15
ID      AAW36797 standard; peptide; 724 AA.
XX      AAW36797;
AC      AAW36797;
XX      23-APR-1998 (first entry)
DT      23-APR-1998 (first entry)
XX      Novel human gene, designated WMP4.
DE      Novel human gene, designated WMP4.
XX      Peptide recognition unit; YAP WW domain binding protein; WBP-1; WBP-2;
KW      WW domain; cell signalling; growth regulation; cytoskeleton organisation;
KM      targeted drug screening; modulator; WW domain interaction; WMP4.
XX      Homo sapiens.
OS      Homo sapiens.
XX      Key
FH      Location/Qualifiers
FT      misc_difference 1..3
FT      /note= "the nucleotides encoding these amino acids are
FT      not given in the specification"
FT      Domain
FT      140..165
FT      /note= "Claimed (claim 49) WW Domain 1"
FT      252..277
FT      /note= "Claimed (claim 49) WW domain 3"
FT      303..328
FT      /note= "Claimed (claim 49) WW domain 3"
FT      618..724
FT      /note= "Claimed (claim 90) HECT domain"
XX      WO9737223-A1.
XX      09-OCT-1997.
XX      03-APR-1997; 97WO-US005547.
XX      03-APR-1996; 96US-00630916.
XX      (CYTO-) CYTOGEN CORP.
XX      (UYNC-) UNIV NORTH CAROLINA.
XX      Pirozzi G, Kay BK, Fowlkes DM;
XX      WPI; 1997-503234/46.
XX      DR      N-PSDB; AAT95700.
XX      PT      Identifying cell signalling and growth regulatory polypeptides by
XX      reaction with multivalent recognition complex - polypeptides are useful
XX      in targeted drug selection.
XX      Claim 48; Fig 23; 220p; English.
XX      The present sequence represents a novel protein WMP4. The WMP4 gene was
XX      identified and isolated from a cDNA expression library generated from
XX      LNCaP prostate cancer cell line, using peptides AAW3603-64. These
XX      peptide recognition units are based on the sequences of WW domain binding
XX      domains of the alpha and gamma subunits of epithelial sodium channel
XX      protein. The WW domain is a small functional domain found in a large
XX      number of proteins from a variety of species including humans, nematodes
XX      and yeast. Its name is derived from the observation that two tryptophan
XX      residues, one in the amino terminal portion of the WW domain and one in
XX      the carboxyl terminal portion, are conserved. Most proteins containing WW
XX      domains have a function involving cell signalling and growth regulation
XX      or the organisation of the cytoskeleton. Polypeptides containing a WW
XX      domain are identified by treating a multivalent recognition unit complex
XX      that has selective binding affinity for a WW domain, with many
XX      polypeptides and identifying those with selective affinity for the
XX      complex. Proteins containing WW domains are used for targeted drug
XX      screening, i.e. to identify potential modulators of specific WW domain
XX      interactions. The valency of the recognition unit is important in

```

```

CC      determining specificity of interaction with WW domains. In multivalent
CC      form specificity is relaxed, but not lost, so proteins containing WW
CC      domains similar, but not identical, to the sequence of the peptides
CC      target WW can be detected, including new polypeptides
XX      SQ      Sequence 724 AA;
Query Match      33.6%; Score 1355.5; DB 2; Length 724;
Best Local Similarity 43.7%; Pred. No. 7.7e-118;
Matches 280; Conservative 98; Mismatches 200; Indels 63; Gaps 13;
137 QSRDRIGTGQGVDCSRLLFDN---DLDPGWEERTASGRIOYLNHTRTTQWERTRPAS 193
111 RARSSTVTGGEPPTSVAVYHTTGLPGGWEERKDAKGRYYVNHNNRTTWTTRPIQOLA 170
194 ESSSG-----RPLSC---FYDENTPISGTCGACGSSDRLRERVRS 235
171 EDGASGSATNSNNHILIEPQIRRPRLSSFTYLSAPGAGADSVRRVAKDTLSNPSQPQ 230
236 QHRNRYMSRTHLHPEDLPEGYEQTTOGQGVYFLHTQGVSTWHDPRVP-----RDLN 290
231 PEPYNSPKRQHKVTSFLPPGWEKRIAPNGSPFFIDHTKTTWEDPLKFPVHMRKTS 290
291 INCELGFLPPGWEIRNTATGRVYFVDHNNRTTQFTPRLSANLHLVLRONQLKDQOOQ 350
291 INPNDLGFLPPGWEIRHLDGRFTFYIDNSKITQWEDPRL-----QN----- 332
351 QVSLCPDPTDECLATVPR--YKRDLVQKLIIRBELSQOQPOAGCRILEVSRFEESYR 408
333 -----PAITGPAPYGRERKQKDYRKKLTKKPADIPIINFEMKLNHNHNFEESSYR 382
409 QVMKRPDLWK-RIMKFRGEESGLDYGVARBNLYLSHEMLNPYYGLFOYSRDIYTL 467
383 RIMSVKRPDVLKARLWIEFESEKGLDYGVARBNFLSKEMFNYYGLFEYSATDYYTL 442
468 QINPDAV-NPEHLYEFHVGIRMGMAVFHGHIYDGGFTLPFYKQLGKSTLDDMLVD 526
443 QINPMSGCNEPDHLSYFTFGKVAGLAVFHKLLDGGFIIRFYKMLGKQITLNDMSVD 502
527 PDLNSLVMLLENDITGLDHTFCVENNAVGEIIOHEIKPNGKSI PVNEENKXKVALYV 586
503 SEYNSLKWLENDPT-ELDLMFCIDEENFQOTYQVLDKPGSEIMVTNEKREYIDLVI 561
587 NWRPLRGLEAQLQKGFNEVIVQHLLKTPDELELIICGKIVNPKNTRLK-- 644
562 QRFVNRVQKONAFLEBGFTELPIDLIKTFDENLEBLMCGLDVDVDMRQHSIYKNG 621
645 HCTPDSNIVKMFKAVERFEERRARLLQFVTSGRVPLQGFKALQGAAGRLLFTIHOID 704
622 YC-PHNPVIOQFWKAVLLMDAERIRLLQFVTGSRVPMNGFALYGSNGQLFTISQWG 680
705 ACTNNLPRAHTCFNRIDIPYESTYEKYLEKLLTIAETCGF 745
681 S-PEKLPRAHTCFNRLLDPFYETFEDELKELMLAVENAGSF 720

```

Search completed: September 21, 2004, 07:43:44
Job time : 98.5466 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OK protein - protein search, using sw model

Run on: September 21, 2004, 07:28:49 ; Search time 89.4534 Seconds
(without alignments)
2283.667 Million cell updates/sec

Title: US-10-009-945-2

Percent score: 3884
Sequence: 1 GGSISIKRLTVLCAKRLAKK.....EKLYEKLTAVERTCGFAVE 723

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database: A_Geneseq_29Jan04:*

1: Geneseq1980s:.*
2: Geneseq1990s:.*
3: Geneseq2000s:.*
4: Geneseq2001s:.*
5: Geneseq2002s:.*
6: Geneseq2003as:.*
7: Geneseq2003bs:.*
8: Geneseq2004s:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3884	100.0	723	4 AAB31476	Aab31476 Amino act
2	3878	99.8	722	6 AAE32724	Aae32724 E3 ubiqui
3	3869	99.6	722	4 AAM79861	Aam79861 Human pro
4	3001	77.3	748	2 AAM13384	Aam13384 Human pro
5	3001	77.3	748	4 AAB31477	Aab31477 Amino act
6	2996	77.1	804	4 AAU19610	Aau19610 Human dia
7	2236	57.6	514	4 AAU87301	Aau87301 Novel cen
8	2069	53.3	1035	4 AAB61120	Aab61120 Drosophi
9	1614.5	41.6	335	4 AAM78877	Aam78877 Human pro
10	1461	37.6	766	5 AAM13385	Aam13385 Human pro
11	1436.5	37.0	832	5 AAM73459	Aam73459 Candida a
12	1345	34.6	869	6 AABU26104	Aabu26104 Aspergill
13	1318	33.9	854	2 AAY30949	Aay30949 Murine E3
14	1301	33.5	852	2 AAY30948	Aay30948 Human E3
15	1288.5	33.2	903	6 ABR64210	Abr64210 Angiogene
16	1259	32.4	927	2 AAY25170	Aay25170 Human KTA
17	1259	32.4	927	6 AAE32725	Aae32725 Human NED
18	1249.5	32.2	911	7 ADB75455	Adb75455 Prostate
19	1245.5	32.1	739	4 AABG16477	Aabg16477 Novel hum
20	1245.5	32.1	739	4 AAG68173	Aag68173 Atrophin-
21	1245.5	32.1	739	6 AAE32720	Aae32720 Atrophin-
22	1245.5	32.1	739	6 ABR41097	Abr41097 Human atr
23	1245.5	32.1	739	7 ADB98717	Adb98717 Human atr
24	1245.5	32.1	739	7 ADB82513	Adb82513 Human pro
25	1245.5	32.1	752	8 ADE77070	Ade77070 Human pro

26	1245	32.1	975	2 AAW93167	Aaw93167 Human ZGC
27	1239	31.9	922	4 AAB50049	Aab50049 Human hom
28	1239	31.9	922	5 AAB05708	Aab05708 Human sig
29	1239	31.9	949	4 ABB59631	Abb59631 Drosophi
30	1239	31.9	949	4 ABB59856	Abb59856 Drosophi
31	1228.5	31.6	923	4 AABG01080	Aabg01080 Novel hum
32	1228.5	31.6	995	6 AAE32719	Aae32719 KIAA0439
33	1228.5	31.6	995	6 AAE32719	Aae32719 Prostate
34	1228.5	31.6	995	7 ADB75451	Adb75451 Prostate
35	1228.5	31.6	995	7 ADB75451	Adb75451 Prostate
36	1227	31.6	724	2 ADD45762	Add45762 Human Pro
37	1227	31.6	725	2 AAB99320	Aab99320 Novel hum
38	1226.5	31.6	759	6 AABU2504	Aabu2504 Aspergill
39	1226	31.6	854	7 ADB75453	Adb75453 Prostate
40	1225	31.5	870	4 AAE05495	Aae05495 Human ubi
41	1225	31.5	870	6 AAE32722	Aae32722 Nedd-4-1
42	1225	31.5	906	7 ADB849242	Adb849242 Novel hum
43	1223	31.5	834	5 AAU77715	Aau77715 Dissected
44	1223	31.5	854	6 ABB58332	Abb58332 Human cel
45	1212	31.2	474	2 AAY10943	Aay10943 Human act

ALIGNMENTS

RESULT 1
ID AAB31476 standard; protein; 723 AA.

AC AAB31476;
DT 20-APR-2001 (first entry)

DE Amino acid sequence of a human Smurf1 polypeptide.

KW Smurf1; Smurf2; Smad signal transduction; bone morphogenic protein; BMP;
KW transforming growth factor-beta; human; TGF-beta; Chondrogenesis;
KW osteogenesis; blood differentiation; cartilage formation; hair growth;
KW neural tube patterning; retinal development; heart induction;
KW morphogenesis; tooth formation; gamete formation.

CS Homo sapiens.

XX WO200077168-A2.

XX 21-DEC-2000.

XX 12-JUN-2000; 2000WO-US016250.

XX 11-JUN-1999; 99US-0138969P.

XX (UNYX) UNIV NEW YORK STATE RES FOUND.
XX (HSCR-) HSC RES & DEV LP.

XX Thomsen GH, Wraan J;
XX WPT, 2001-071267/08.

XX N-PDB; AAF24852.

PT Novel isolated Smurf protein useful for inhibiting bone morphogenic
PT protein or tumor growth factor-beta activation pathway, for treating
PT cancer and to block osteogenesis, hair growth, tooth formation.

PS Claim 6; Fig 10; 107pp; English.

CC The present sequence represents a human Smurf1 polypeptide. The
CC specification also describes a Smurf1 polypeptide. Smurf polypeptides are
CC negative regulators of Smad signal transduction, and antagonists of bone
CC morphogenic protein (BMP) or transforming growth factor-beta (TGF-beta)
CC signaling pathway. Expression of Smurf1 in a cell is useful for
CC inhibiting a BMP or TGF-beta activation pathway in a cell. Smurf
CC polypeptides are useful for blocking chondrogenesis, osteogenesis, blood
CC differentiation, cartilage formation, neural tube patterning, retinal

CC development, heart induction and morphogenesis, hair growth, tooth
 CC formation, gamete formation and a wide variety of tissue and organ
 CC formation processes, and hinder the regeneration, growth, maintenance,
 CC etc., of bone and other tissues that are dependent on the BMP pathway.
 CC The polypeptide is useful for screening for various drugs and/or
 CC antibodies that can either enhance the BMP pathway, or inhibit it
 CC

Sequence 723 AA;

Query Match 100.0%; Score 3884; DB 4; Length 723;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 723; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GSSIKIRLTVLCANLAKKDFRLLPDPFAKIVVDSGQCHSTVTKTLDPKKNQHYDL 60
 DB 1 GSSIKIRLTVLCANLAKKDFRLLPDPFAKIVVDSGQCHSTVTKTLDPKKNQHYDL 60
 QY 61 YVGTDSITISVNMHKKIHKQAGAGFLGCVRLSNALSRLKDTGYQRLDCKLNPSDTDA 120
 DB 61 YVGTDSITISVNMHKKIHKQAGAGFLGCVRLSNALSRLKDTGYQRLDCKLNPSDTDA 120
 QY 121 VVGQIVLSLQTRDRIGTGSVYVDCRGLLENNGTYEDSGPRPLSCMEBPAPYTDSTGA 180
 DB 121 VVGQIVLSLQTRDRIGTGSVYVDCRGLLENNGTYEDSGPRPLSCMEBPAPYTDSTGA 180
 QY 181 AAGGNCRFVESPQDQRLQARLNPDPVRSGLQTPQNRPHGQSPPLPSGEQRTTVQG 240
 DB 181 AAGGNCRFVESPQDQRLQARLNPDPVRSGLQTPQNRPHGQSPPLPSGEQRTTVQG 240
 QY 241 QVFLHTQGVSTWHDPRIPRDLNSVNCDELGPIPPGMEVRSVSGRIYVDHNNRTTOFT 300
 DB 241 QVFLHTQGVSTWHDPRIPRDLNSVNCDELGPIPPGMEVRSVSGRIYVDHNNRTTOFT 300
 QY 301 TDPRLHIMNHQCOLKEPSQPLPLPSSEGSLEDEELPAQRYERDLVQKLKLRHELSIQP 360
 DB 301 TDPRLHIMNHQCOLKEPSQPLPLPSSEGSLEDEELPAQRYERDLVQKLKLRHELSIQP 360
 QY 361 QAGHCRLEVSREELFEESYROIIMKRPDLKRLMKVFRGEGLDYGVAREVLYLICHE 420
 DB 361 QAGHCRLEVSREELFEESYROIIMKRPDLKRLMKVFRGEGLDYGVAREVLYLICHE 420
 QY 421 MLNYYGLFQYSTDNIVMLQINPDSSINPDLSYFHFVGRIMGLAVFHGYINGGFTVPF 480
 DB 421 MLNYYGLFQYSTDNIVMLQINPDSSINPDLSYFHFVGRIMGLAVFHGYINGGFTVPF 480
 QY 481 YKQLISPRQLQSLDESVDPELHKSGLVILENDITPVLDHTFCVEHNAFGRIILQHELKPN 540
 DB 481 YKQLISPRQLQSLDESVDPELHKSGLVILENDITPVLDHTFCVEHNAFGRIILQHELKPN 540
 QY 541 RNVPTEENKKEVYRLVYVNRPMRGIEAOFALQKGFNELIPQHLKPEQKLELITGG 600
 DB 541 RNVPTEENKKEVYRLVYVNRPMRGIEAOFALQKGFNELIPQHLKPEQKLELITGG 600
 QY 601 LDKTDLNDKMSNRELKHCVADSNIVRFWQAVETFEDEERARLLQFVTSRVPLOGFPA 660
 DB 601 LDKTDLNDKMSNRELKHCVADSNIVRFWQAVETFEDEERARLLQFVTSRVPLOGFPA 660
 QY 661 LQSTGAGARLFTIHIDANTONTLPKATCCNRIDIPPESEYKLYEKLTAVEETCGF 720
 DB 661 LQSTGAGARLFTIHIDANTONTLPKATCCNRIDIPPESEYKLYEKLTAVEETCGF 720
 QY 721 AVE 723
 DB 721 AVE 723

RESULT 2

AAE32724 ID AAE32724 standard; protein; 722 AA.
 XX AAE32724;
 AC
 XX
 DT 24-MAR-2003 (first entry)

XX E3 ubiquitin ligase SMURP1 protein.
 DE
 XX Viral infection; lymphosarcoma; human immunodeficiency virus; hepatitis;
 KM polioyelitis; HIV; measles; protein therapy; E3 ubiquitin ligase;
 XX enzyme.
 KM

OS Unidentified.

XX WO200290549-A2.

XX 14-NOV-2002.

XX 12-MAR-2002; 2002WO-1B002106.

XX 12-MAR-2001; 2001US-0275224P.

XX 31-JUL-2001; 2001US-0308958P.

XX 07-DEC-2001; 2001US-0340170P.

XX (PROT-) PROTEOLOGICS LTD.

XX Greener T, Moskowitz H, Reiss Y, Alroy I;

XX MPI; 2003-111976/10.

XX N-PSDB; AAD50458.

XX New protein complex comprising HECT-RCCL, viral maturation scaffolding
 PT protein (VNSP), and/or HIV gag protein, useful for treating viral
 PT infections, such as lymphosarcoma, HIV, hepatitis, polioyelitis, measles,
 PT or Ebola.

XX Disclosure; Fig 14; 150pp; English.

XX The invention relates to a method for modulation of viral maturation. The
 CC invention also provides an isolated protein complex comprising a HECT-
 CC RCCL selected from HECT-WM, HECT-RCCL, Gag protein, Gag late domain, P13,
 CC actin, myosin, Hsp60, Hsp70, Hsp90, STAM1, STAM2A, STAM2B, VHS-UM,
 CC GTPase, E2 enzyme, tsgl01, cullin, HECT1, HECT2, HECT3, Ned4-1-like
 CC protein or clathrin. The complexes, proteins, antibodies and methods are
 CC useful for treating viral infections, such as lymphosarcoma, human
 CC immunodeficiency virus (HIV), hepatitis, polioyelitis, measles, or Ebola
 CC and for inhibiting budding in a subject. They are also useful in
 CC diagnostic assays for determining whether a cell is infected with a virus
 CC and for characterizing the nature, progression and/or infectivity of the
 CC infection. The invention is also useful in protein therapy. The present
 CC sequence is E3 ubiquitin ligase SMURP1 protein used to illustrate the
 CC method of the invention

XX Sequence 722 AA;

Query Match 99.8%; Score 3878; DB 6; Length 722;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 722; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 GSSIKIRLTVLCANLAKKDFRLLPDPFAKIVVDSGQCHSTVTKTLDPKKNQHYDL 61
 DB 1 GSSIKIRLTVLCANLAKKDFRLLPDPFAKIVVDSGQCHSTVTKTLDPKKNQHYDL 60
 QY 62 YVGTDSITISVNMHKKIHKQAGAGFLGCVRLSNALSRLKDTGYQRLDCKLNPSDTDA 121
 DB 61 YVGTDSITISVNMHKKIHKQAGAGFLGCVRLSNALSRLKDTGYQRLDCKLNPSDTDA 120
 QY 122 VVGQIVLSLQTRDRIGTGSVYVDCRGLLENNGTYEDSGPRPLSCMEBPAPYTDSTGA 181
 DB 121 VVGQIVLSLQTRDRIGTGSVYVDCRGLLENNGTYEDSGPRPLSCMEBPAPYTDSTGA 180
 QY 182 AAGGNCRFVESPQDQRLQARLNPDPVRSGLQTPQNRPHGQSPPLPSGEQRTTVQG 241
 DB 181 AAGGNCRFVESPQDQRLQARLNPDPVRSGLQTPQNRPHGQSPPLPSGEQRTTVQG 240
 QY 242 VYFLHTQGVSTWHDPRIPRDLNSVNCDELGPIPPGMEVRSVSGRIYVDHNNRTTOFT 301
 DB 241 VYFLHTQGVSTWHDPRIPRDLNSVNCDELGPIPPGMEVRSVSGRIYVDHNNRTTOFT 300

QY 302 DPLRHHIMNHQCOLKEPSQPLPLPSEGSLEDEELPAQRYERDLYOKLKYLRHELSLOQPQ 361
 DB 301 DPLRHHIMNHQCOLKEPSQPLPLPSEGSLEDEELPAQRYERDLYOKLKYLRHELSLOQPQ 360
 QY 362 AGHCRIEVSREELFEESYRQIMMRPKDKKRLMVKFGEGLDYGVAREWLYLLCHEM 421
 DB 361 AGHCRIEVSREELFEESYRQIMMRPKDKKRLMVKFGEGLDYGVAREWLYLLCHEM 420
 QY 422 LNPYYGLFQYSTNTIYMLQINPDSSINPHLSYFHFVGRIMGLAVFHGYINGGFTVPY 481
 DB 421 LNPYYGLFQYSTNTIYMLQINPDSSINPHLSYFHFVGRIMGLAVFHGYINGGFTVPY 480
 QY 482 KOLLGKPIQLSDLESVDPELHKSILWILLENDITPVLDFHFCVENHAFGRILLOHELKPNR 541
 DB 481 KOLLGKPIQLSDLESVDPELHKSILWILLENDITPVLDFHFCVENHAFGRILLOHELKPNR 540
 QY 542 NVPVTEENKKEYVRLYVNMRFMRGIEAQLALOKGFNELIPQHLKPFQCKELELLIGGL 601
 DB 541 NVPVTEENKKEYVRLYVNMRFMRGIEAQLALOKGFNELIPQHLKPFQCKELELLIGGL 600
 QY 602 DKIDLNDMKSNTRLKHCVADSNIVRMFWQAVETFEDEERRARLLQVTSSTRVPLQGFAL 661
 DB 601 DKIDLNDMKSNTRLKHCVADSNIVRMFWQAVETFEDEERRARLLQVTSSTRVPLQGFAL 660
 QY 662 QGSTGAAGPRLFTIHLIDANTDNLPRKATCFNRIDIPPESEYKLYEKLITAVEETCGFA 721
 DB 661 QGSTGAAGPRLFTIHLIDANTDNLPRKATCFNRIDIPPESEYKLYEKLITAVEETCGFA 720
 QY 722 VE 723
 DB 721 VE 722

RESULT 3
 AAM79861
 ID AAM79861 standard; protein: 722 AA.
 XX
 AC AAM79861;
 XX
 DT 06-NOV-2001 (first entry)
 XX
 DE Human protein SEQ ID NO 3507.
 XX
 KM Human; cytokine; cell proliferation; cell differentiation; gene therapy;
 KM vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
 KM tissue growth factor; immunomodulatory; cancer; leukaemia;
 KM nervous system disorder; arthritis; inflammation.
 XX
 OS Homo sapiens.
 XX
 PN WC200157190-A2.
 XX
 PD 09-AUG-2001.
 XX
 PE 05-FEB-2001; 2001WO-US004098.
 XX
 PF 03-FEB-2000; 2000US-00496914.
 PR 27-APR-2000; 2000US-00560875.
 PR 20-JUN-2000; 2000US-00598075.
 PR 19-JUL-2000; 2000US-00620325.
 PR 01-SEP-2000; 2000US-00654936.
 PR 15-SEP-2000; 2000US-00663561.
 PR 20-OCT-2000; 2000US-00693325.
 PR 30-NOV-2000; 2000US-00728422.
 XX
 XX (HYSE-) HYSEQ INC.
 PA Tang YT, Liu C, Dimaac RT, Asundi V, Zhou P, Xu C, Cao Y;
 PI Ma Y, Zhao QA, Wang J, Zhang J, Ren F, Chen R, Wang ZW;
 PI Xue AJ, Yang Y, Wejhtman T, Goodrich R;
 XX WPI; 2001-476283/51.

DR N-PSDB; AAK52994.
 XX Nucleic acids encoding polypeptides with cytokine-like activities, useful
 PT in diagnosis and gene therapy.
 PT
 XX
 XX Claim 20; Page 362-363; 6221pp; English.
 PS
 XX
 CC The invention relates to polynucleotides (AAK51456-AAK53435) and the
 CC encoded polypeptides (AAM78323-AAK80302) that exhibit activity elating to
 CC cytokine, cell proliferation or cell differentiation or which may induce
 CC production of other cytokines in other cell populations. The
 CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
 CC peptide therapy. The polypeptides have various cytokine-like activities,
 CC e.g. stem cell growth factor activity, haematopoiesis regulating
 CC activity, tissue growth factor activity, immunomodulatory activity and
 CC activity/inhibin activity and may be useful in the diagnosis and/or
 CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
 CC inflammation. Note: Records for SEQ ID NO 2110 (AAK52581), 2111
 CC (AAK52582) and 3666 (AAM80020) are omitted as the relevant pages from the
 CC sequence listing were missing at the time of publication
 CC
 XX
 SQ Sequence 722 AA;
 Query Match 99.6%; Score 3869; DB 4; Length 722;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 721; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 2 GSSIKIRLTVLCAKMLAKDFFRLDPFAKIVVDGSGGCHSTDTYKNTLDPKNOHIDLY 61
 DB 1 GSSIKIRLTVLCAKMLAKDFFRLDPFAKIVVDGSGGCHSTDTYKNTLDPKNOHIDLY 60
 QY 62 VKKTDSTITISWNHKKIKKKGAGTGCVRLLSNASISLKDGYORLDCLNPSDTPAY 121
 DB 61 VKKTDSTITISWNHKKIKKKGAGTGCVRLLSNASISLKDGYORLDCLNPSDTPAY 120
 QY 122 RQIIVASLQTRDRIGTGSVYDCRGLNENGTVEYDSGPGSPPLSCFMEBPAPYTDSTGAA 181
 DB 121 RQIIVASLQTRDRIGTGSVYDCRGLNENGTVEYDSGPGSPPLSCFMEBPAPYTDSTGAA 180
 QY 182 AGGNCGRFVESPSQORLQAOQLRNPVDRGSLQTPQNRPHGHOSPELDEGEORRTVGO 241
 DB 181 AGGNCGRFVESPSQORLQAOQLRNPVDRGSLQTPQNRPHGHOSPELDEGEORRTVGO 240
 QY 242 VFELHTQGVSTWHPRLPRDLSVNCDELGPLPFGMEVSTVSGRIFVCHNNRTTOFT 301
 DB 241 VFELHTQGVSTWHPRLPRDLSVNCDELGPLPFGMEVSTVSGRIFVCHNNRTTOFT 300
 QY 302 DPLRHHIMNHQCOLKEPSQPLPLPSEGSLEDEELPAQRYERDLYOKLKYLRHELSLOQPQ 361
 DB 301 DPLRHHIMNHQCOLKEPSQPLPLPSEGSLEDEELPAQRYERDLYOKLKYLRHELSLOQPQ 360
 QY 362 AGHCRIEVSREELFEESYRQIMMRPKDKKRLMVKFGEGLDYGVAREWLYLLCHEM 421
 DB 361 AGHCRIEVSREELFEESYRQIMMRPKDKKRLMVKFGEGLDYGVAREWLYLLCHEM 420
 QY 422 LNPYYGLFQYSTNTIYMLQINPDSSINPHLSYFHFVGRIMGLAVFHGYINGGFTVPY 481
 DB 421 LNPYYGLFQYSTNTIYMLQINPDSSINPHLSYFHFVGRIMGLAVFHGYINGGFTVPY 480
 QY 482 KOLLGKPIQLSDLESVDPELHKSILWILLENDITPVLDFHFCVENHAFGRILLOHELKPNR 541
 DB 481 KOLLGKPIQLSDLESVDPELHKSILWILLENDITPVLDFHFCVENHAFGRILLOHELKPNR 540
 QY 542 NVPVTEENKKEYVRLYVNMRFMRGIEAQLALOKGFNELIPQHLKPFQCKELELLIGGL 601
 DB 541 NVPVTEENKKEYVRLYVNMRFMRGIEAQLALOKGFNELIPQHLKPFQCKELELLIGGL 600
 QY 602 DKIDLNDMKSNTRLKHCVADSNIVRMFWQAVETFEDEERRARLLQVTSSTRVPLQGFAL 661
 DB 601 DKIDLNDMKSNTRLKHCVADSNIVRMFWQAVETFEDEERRARLLQVTSSTRVPLQGFAL 660
 QY 662 QGSTGAAGPRLFTIHLIDANTDNLPRKATCFNRIDIPPESEYKLYEKLITAVEETCGFA 721

Db 661 QGSGAAGRLFTLHIDANTNDNRKATCFNRIDIPYSEYKYLTAVEETCGFA 720
QY 722 VE 723
Db 721 VE 722

RESULT 4
AAW13384
ID AAW13384 standard; protein; 748 AA.
XX
AC AAW13384;
XX

DT 10-JUN-1997 (first entry)
XX
DE Human protein ubiquitin ligase pub1.

XX Protein ubiquitin ligase; pub1; cdc25 phosphatase; CDK kinase; p53;
XX cell cycle; transgenic animal.

OS Homo sapiens.

XX WO9712962-A1.

XX 10-APR-1997.

XX 04-OCT-1996; 96WO-US015930.

XX 04-OCT-1996; 95US-00539205.

XX (COLD-) COLD SPRING HARBOR LAB.

XX Beach D, Caligiuri M, Nefsky B;

XX WPI; 1997-226206/20.

XX N-PSDB; AAT47040.

XX Human and Saccharomyces pombe protein ubiquitin ligase(s) - involved in
XX cdc25 phosphatase and p53 ubiquitination, and regulate cell growth and
XX proliferation.

XX Claim 1; Page 74-77; 108pp; English.

XX Human protein ubiquitin ligases pub1 (AAW13384), pub2 (AAW13385) and pub3
XX (AAW13386) are homologues of fission yeast pub1 (AAW13387) and were
XX identified from cDNA clones (AAT47040-42) obt. e.g. from a keratinocyte
XX cDNA library. Pub polypeptides can be produced in transfected host cells.
XX They can control the steady state level of cdc25 phosphatase, the degree
XX of CDK kinase (e.g. cdc2) dephosphorylation and the steady state level of
XX p53 (controlling the degree of cell cycle regulation of p53). They can
XX regulate cell or tissue differentiation, or cell growth or proliferation
XX by affecting other proteins, can be a specific (ant)agonist of wild-type
XX protein function and may be used as immunogens to elicit a specific
XX immune residue

XX Sequence 748 AA;

XX Query Match 77.3%; Score 3001; DB 2; Length 748;

XX Best Local Similarity 74.6%; Pred. No. 1.5e-285;

XX Matches 565; Conservative 63; Mismatches 71; Indels 58; Gaps 8;

QY 5 IKILTYCACNLAKKPPFLPDPFAKIVVDGSGCSTDTVKTLDPKKNQHYDYVKK 64

Db 12 VKELTYCACNLAKKPPFLPDPFAKIVVDGSGCSTDTVKTLDPKKNQHYDYVKK 71

QY 65 TDSITISVYNNHKKIHKQAGFJGCVRLLSNAISRLKDTGYORLDCIKLNPSTDAVGG 124

Db 72 SDSVTISVYNNHKKIHKQAGFJGCVRLLSNAISRLKDTGYORLDCIKLNPSTDAVGG 131

QY 125 IVVSLQTRDRIGTGGSVYDCKGLLEN-----GTIV----- 155

Db 132 IVVSLQTRDRIGTGGSVYDCKGLLEN-----GTIV----- 191

QY 156 --EDSGPRPLSCFMEBPAPYTDSTGAAGGNCRFVSPSODRLQARLNPVGRSL 213
Db 192 ASEYSSGPRPLSCFMEBPAPYTDSTGAAGGNCRFVSPSODRLQARLNPVGRSL 242

QY 214 QTPQNPAGHQBELPEGYEQRTTVQGVYFLHTQVSTWMDPRLPDLSVNDDELGP 273

Db 243 ---SRTHLTPPDLPEGYEQRTTVQGVYFLHTQVSTWMDPRLPDLSVNDDELGP 298

QY 274 LPPGWEVSTVSGRIFYVDHNNRTTQFTDPR---LHHIMNQCLKEPSPQLPLPSBGS 329

Db 299 LPPGWEVSTVSGRIFYVDHNNRTTQFTDPR---LHHIMNQCLKEPSPQLPLPSBGS 354

QY 330 L---EDSELPAGRYERDLVQKLVTRHELSLQOPQAGCRLEVSREEIFEESYRQMKR 386

Db 355 LCPDDTECLTVPRYKRDVQKLVTRHELSLQOPQAGCRLEVSREEIFEESYRQMKR 414

QY 387 PKDLKRLVWFRGEGGLDYGVAREMWYLCEKMLNPYYGLFOYSTNINWLOJNPSS 446

Db 415 PKDLKRLVWFRGEGGLDYGVAREMWYLCEKMLNPYYGLFOYSTNINWLOJNPSS 474

QY 447 INPDHLYFHYVGRINGLAVFHHYINGFTVPPYKQLGKPIQLSDLESVDPELHSLV 506

Db 475 VNPHELSYFHYVGRINGLAVFHHYINGFTVPPYKQLGKPIQLSDLESVDPELHSLV 534

QY 507 WLENDITPVLDHTCVENHNAFGRILOHKLKNGRNVPTENKKEAYRLVYNNRPMRG 566

Db 535 WLENDITPVLDHTCVENHNAFGRILOHKLKNGRNVPTENKKEAYRLVYNNRPMRG 594

QY 567 EAQFLALQGFNELIPOHLKPPDQKELELIIGSLDKIDLNDWKSNTLKHCVADSNIVR 626

Db 595 EAQFLALQGFNELIPOHLKPPDQKELELIIGSLDKIDLNDWKSNTLKHCVADSNIVR 654

QY 627 WFMQAVETFEERRRRLQFVTSRVPLQGRKALQSTGAGPRLFTIHLIDANTDLP 686

Db 655 WFMQAVETFEERRRRLQFVTSRVPLQGRKALQSTGAGPRLFTIHLIDANTDLP 711

QY 687 KATCFNRIDIPYSEYKYLTAVEETCGFAVE 723

Db 712 KATCFNRIDIPYSEYKYLTAVEETCGFAVE 748

RESULT 5

AAAB31477 standard; protein; 748 AA.

AAAB31477;

DT 20-APR-2001 (first entry)

XX Amino acid sequence of a human Smurf2 polypeptide.

XX Smurf1; Smurf2; Smad signal transduction; bone morphogenic protein; BMP;

XX transforming growth factor-beta; human; TGF-beta; Chondrogenesis;

XX osteogenesis; blood differentiation; cartilage formation; hair growth;

XX neural tube patterning; retinal development; heart induction;

XX morphogenesis; tooth formation; gamete formation.

XX Homo sapiens.

XX WO200071168-A2.

XX 21-DEC-2000.

XX 12-JUN-2000; 2000WO-US016250.

XX 11-JUN-1999; 99US-0138969P.

XX (UNYV) UNIV NEW YORK STATE RES FOUND.

XX (HSCR-) HSC RES & DEV LP.

XX Thomsen GH, Wraza J;

XX WPI; 2001-071267/08.

Spilant

DR. N-PSDB; AAF24853.

PT Novel isolated Smurf protein useful for inhibiting bone morphogenic protein or tumor growth factor-beta activation pathway, for treating cancer and to block osteogenesis, hair growth, tooth formation.

PS Claim 10; Fig 12; 107pp; English.

XX The present sequence represents a human Smurf2 polypeptide. The CC specification also describes a Smurf1 polypeptide. Smurf polypeptides are CC negative regulators of Smad signal transduction, and antagonists of bone CC morphogenic protein (BMP) or transforming growth factor-beta (TGF-beta) CC signaling pathway. Expression of Smurf1 in a cell is useful for CC inhibiting a BMP or TGF-beta activation pathway in a cell. Smurf CC polypeptides are useful for blocking chondrogenesis, osteogenesis, blood CC differentiation, cartilage formation, neural tube patterning, retinal CC development, heart induction and morphogenesis, hair growth, tooth CC formation, gamete formation and a wide variety of tissue and organ CC formation, processes, and hinder the regeneration, growth, maintenance, CC etc., of bone and other tissues that are dependent on the BMP pathway. CC The polypeptide is useful for screening for various drugs and/or CC antibodies that can either enhance the BMP pathway, or inhibit it CC

Sequence 748 AA.

Query Match 77.3%; Score 3001; DB 4; Length 748;

Best Local Similarity 74.6%; Pred. No. 1,5e-285; Matches 565; Conservative 63; Mismatches 71; Indels 58; Gaps 8;

QY 5 IKTITLTKAKNLAKKDFFRLPDPKAIYVDGSGGCHSTDTYNTKADPKMKNHYDLYGK 64
 DB 12 VKRLTVLCKAKNLYKXDFRLPDPKAIYVDGSGGCHSTDTYNTKADPKMKNHYDLYGK 71
 QY 65 TDSITLTVNHNKHKIKKQAGAGLGCYVLLSNALSLKDTGYOBLDCKLNPSTDAVRQ 124
 DB 72 SDSVTLSVNMHNKHKIKKQAGAGLGCYVLLSNALSLKDTGYOBLDCKLNPSTDAVRQ 131
 QY 125 IVVSLQTRDRIGTGGSYVDGRGLIENE-----GTVY----- 155
 DB 132 IVVSLQTRDRIGTGGSYVDGRGLIENE-----GTVY----- 155
 QY 132 IVVSLQTRDRIGTGGSYVDGRGLIENE-----GTVY----- 155
 DB 132 IVVSLQTRDRIGTGGSYVDGRGLIENE-----GTVY----- 155
 QY 156 --EDSGGRPLSCFMEBPAPYDSTGAAGGNCRFVESPDQRLQACRLNPDVAGSL 213
 DB 192 ASEVSSSGRPLSCFMEBPAPYDSTGAAGGNCRFVESPDQRLQACRLNPDVAGSL 213
 QY 214 QTPONRPHGHSPLPFGYEGRTTVGQVYFLHTQTGVSFWHPRIKRLDANSVNCDELGP 273
 DB 243 ---SRTHLTPPDLPGEGEORTTQGGVYFLHTQTGVSFWHPRIKRLDANSVNCDELGP 273
 QY 274 LPPGMEVRSYSGRIYFDHNHRTTQFTDPR---LHINMHCQLKEPQPLPLPSEGS 329
 DB 299 LPPGMEVRSYSGRIYFDHNHRTTQFTDPR---LHINMHCQLKEPQPLPLPSEGS 329
 QY 330 L---EDELPAORYERDLYQKLYRLHLSLQOPQACHCRILEVSRBEIIFBESRQIMKMR 366
 DB 355 LCPDTECLTVPRYKRDLYQKLYRLHLSLQOPQACHCRILEVSRBEIIFBESRQIMKMR 366
 QY 387 PKDLKRLMKVPRBEGADYGVAAREMYLLCHEMLNPPYGLFQYSIDNTYMLQINPDS 446
 DB 415 PKDLKRLMKVPRBEGADYGVAAREMYLLCHEMLNPPYGLFQYSIDNTYMLQINPDS 446
 QY 447 INPDHLSYFHFVGRIMGLAVFHGYINGFTVPFYKQLGKPLQSLDESVPDLKSLV 506
 DB 475 VNPBLSYFHFVGRIMGLAVFHGYINGFTVPFYKQLGKPLQSLDESVPDLKSLV 506
 QY 507 WILENDITPVLDHFTFCVCHNAFGRILCHELKNQANVPVEENKKEYRLVVMRPMGI 566
 DB 535 WILENDITPVLDHFTFCVCHNAFGRILCHELKNQANVPVEENKKEYRLVVMRPMGI 566
 QY 567 BAQFLALQKGFNELIPQHLKRFQDKELDLIIGGLDKIDLNDKSNTRLKHCIVADSNIVR 626
 DB 595 BAQFLALQKGFNELIPQHLKRFQDKELDLIIGGLDKIDLNDKSNTRLKHCIVADSNIVR 626

QY 627 WFWQAVETPDEERARLILQVTSSTRVPLQGFALQSGSTGAGRPFTLHLDANTDNP 686
 DB 655 WFWQAVETPDEERARLILQVTSSTRVPLQGFALQSGSTGAGRPFTLHLDANTDNP 711
 QY 687 KAHTCFNRIDIPYESYEKLYEKLTAVEETCGFAVE 723
 DB 712 KAHTCFNRIDIPYESYEKLYEKLTAVEETCGFAVE 748

RESULT 6

ID AAU19610 standard; protein; 804 AA.

AC AAU19610;

DT 04-DEC-2001 (first entry)

DE Human diagnostic and therapeutic polypeptide (DITAP) #196.

KM Human; receptor; diagnostic; therapeutic; gene therapy; vaccine;

KW cell proliferative disorder; Crohn's disease; lymphoma; leukemia;

KW acquired immune deficiency syndrome; AIDS; autoimmune disorder;

OS Homo sapiens.

PN WO200162927-A2.

PD 30-AUG-2001.

XX 21-FEB-2001; 2001WO-US006059.

XX 24-FEB-2000; 2000US-0184693P.

XX 24-FEB-2000; 2000US-0184697P.

XX 24-FEB-2000; 2000US-0184698P.

XX 24-FEB-2000; 2000US-0184768P.

XX 24-FEB-2000; 2000US-0184769P.

XX 24-FEB-2000; 2000US-0184770P.

XX 24-FEB-2000; 2000US-0184771P.

XX 24-FEB-2000; 2000US-0184772P.

XX 24-FEB-2000; 2000US-0184773P.

XX 24-FEB-2000; 2000US-0184774P.

XX 24-FEB-2000; 2000US-0184776P.

XX 24-FEB-2000; 2000US-0184797P.

XX 24-FEB-2000; 2000US-0184813P.

XX 24-FEB-2000; 2000US-0184813P.

XX 24-FEB-2000; 2000US-0185213P.

XX 24-FEB-2000; 2000US-0185216P.

XX 12-MAY-2000; 2000US-0203785P.

XX 15-MAY-2000; 2000US-0204226P.

XX 16-MAY-2000; 2000US-0204226P.

XX 16-MAY-2000; 2000US-0204821P.

XX 16-MAY-2000; 2000US-0204908P.

XX 16-MAY-2000; 2000US-0205232P.

XX 17-MAY-2000; 2000US-0204815P.

XX 17-MAY-2000; 2000US-0204863P.

XX 17-MAY-2000; 2000US-0205221P.

XX 17-MAY-2000; 2000US-0205285P.

XX 17-MAY-2000; 2000US-0205286P.

XX 17-MAY-2000; 2000US-0205287P.

XX 17-MAY-2000; 2000US-0205323P.

XX 17-MAY-2000; 2000US-0205324P.

XX (INCYTE) INCYTE GENOMICS INC.

XX Panzer SR, Spiro PA, Banville SC, Shah P, Chalup MS, Chang SC;

XX Chen A, D'sa SA, Amey S, Dahl CR, Daniels SE, Dufour GB;

XX Flores V, Fong WT, Greenwalt LB, Hillman JL, Jones AJ, Liu TF;

XX Roseberry AM, Rosen BH, Russo FD, Stockdreher TK, Datto A;

XX Wright RJ, Yap PE, Bradley DL, Bretcher SR, Chen W;

XX Chen HJ, Hodgson DM, Lincoln SE, Jackson S;

XX WPI; 2001-50267/55.
 DR N-PSDB; AAS31181.
 XX
 PT Polynucleotides encoding diagnostic and therapeutic proteins, e.g.
 XX enzymes, hormones and receptors, useful in diagnostics and therapeutics.
 PS Claim 27; Page 512-514; 522pp; English.
 XX
 CC The invention relates to polynucleotides (I) encoding diagnostic and
 CC therapeutic (DTHP) polypeptides (II), which include e.g. enzymes, and
 CC proteins involved in growth and development and receptors. (I) and (II)
 CC may be used in the prevention, diagnosis and treatment of diseases
 CC associated with inappropriate DTHP expression. For example, (I) and (II)
 CC may be used to treat disorders associated with decreased polypeptide
 CC expression by rectifying mutations or deletions in a patient's genome,
 CC that affect the activity of the DTHPs, by expressing inactive proteins
 CC or supplementing the patient's own production of them. (I) and (II) may
 CC be used to treat diseases, for example, cell proliferative disorder,
 CC Crohn's disease, acquired immune deficiency syndrome (AIDS), lymphoma,
 CC leukaemia, autoimmune disorders, and respiratory disorders. Additionally,
 CC (I) may be used to produce the DTHPs, by inserting the nucleic acids
 CC into a host cell and culturing the cell to express the protein. (I) and
 CC its complementary sequences may also be used as DNA probes in diagnostic
 CC assays to detect and quantify the presence of similar nucleic acids in
 CC samples, and therefore which patients may be in need of restorative
 CC therapy. (II) may also be used as antigens in the production of
 CC antibodies against DTHPs and in assays to identify modulators of DTHP
 CC expression and activity. The anti-DTHP antibodies and antagonists may
 CC also be used to down regulate expression and activity. The anti-DTHP
 CC antibodies may also be used as diagnostic agents for detecting the
 CC presence of DTHPs in samples (e.g. by enzyme linked immunosorbent assay
 CC (ELISA)). AAU9415-AAU9625 represent human diagnostic and therapeutic
 CC (DTHP) polypeptides of the invention
 XX
 XX Sequence 804 AA;
 SQ
 Query Match 77.1%; Score 2996; DB 4; Length 804;
 Best Local Similarity 74.5%; Pred. No. 5.4e-285;
 Matches 564; Conservative 63; Mismatches 72; Indels 58; Gaps 8;
 QY 5 IKIRITVLCANLAKDFFRLDPFAKIVVDGSGQCHSTDTVKTLPKMNQHYDLVYG 64
 DB 68 VMLHLTVLCANLAKDFFRLDPFAKIVVDGSGQCHSTDTVKTLPKMNQHYDLVYG 127
 QY 65 TDSITISVMNKKTHKKAGAGLGVRLSNALSLMDTQVRLDCKLNPSDVDANRGQ 124
 DB 128 SDSVTISVMNKKTHKKAGAGLGVRLSNALSLMDTQVRLDCKLNPSDVDANRGQ 167
 QY 125 IVSLQTRDRIGTGGSVVDCRGLDENE-----GTVV----- 155
 DB 188 IVSLQTRDRIGTGGSVVDCRGLDENE-----GTVV----- 155
 QY 156 --EDSGPRRLSCMEBAPYTDSTGAAGGNCRFVESPQDORLQAQLRNPVAGSL 213
 DB 248 ASESISSEPRPLSCVDEVTPISTGNVATCG-----OSSDRLEAERVRKQQRHNVM 298
 QY 214 QTPNRRPHGQSPELPEGEQRTVQGVYFLHTQTVSTWHDRIIPRDNVNCDELGP 273
 DB 299 ----SRTHLTPRLPPEGEQRTVQGVYFLHTQTVSTWHDRIIPRDNVNCDELGP 354
 QY 274 LPPGQVASTVSGRIYFYDHNRTTQTDPR---LHINNHQCLKEPSQPLPDESGS 329
 DB 355 LPPGQVASTVSGRIYFYDHNRTTQTDPR---LHINNHQCLKEPSQPLPDESGS 410
 QY 330 L---EDELPAQRYERDLYOKLKYLRHSLAQQAQCHRIEVSREIEFESRQIMKMR 366
 DB 411 LCPDTECLVYPRKRDLYOKLKYLRHSLAQQAQCHRIEVSREIEFESRQIMKMR 470
 QY 387 FKDIKKRLMYKRGEEGLDYGVARLWYLLCHEMLNPPYGLFOYSTDNVYMLQINPDS 446
 DB 471 FKDIKKRLMYKRGEEGLDYGVARLWYLLCHEMLNPPYGLFOYSTDNVYMLQINPDS 530

QY 447 INPDHLSYFHFVGRINGLAVFHGTYNGGFTVPFVYKQLLGKPIQLSDLESVDPELHKSIV 506
 DB 531 VNPBLSYFHFVGRINGLAVFHGTYNGGFTVPFVYKQLLGKPIQLSDLESVDPELHKSIV 590
 QY 507 WLENDITPVLDHTPFCVEHNAFGRILOHELKPGNRNVPYTEENKKEYVLYVNWRFMRGI 566
 DB 591 WLENDITGVLDHTPFCVEHNAFGRILOHELKPGNRNVPYTEENKKEYVLYVNWRFMRGI 650
 QY 567 EAQFLAQGFENLHQHLKPKDCKELIIGSLDKIDINMKSNTRKHQVAASNIVR 626
 DB 651 EAQFLAQGFENLHQHLKPKDCKELIIGSLDKIDINMKSNTRKHQVAASNIVR 710
 QY 627 MFWQAVETDEERRARLLQFTVGTSTVPLQGFALQSGTGAAPRLFTIHLIDANTDNP 686
 DB 711 MFWKAVETDEERRARLLQFTVGTSTVPLQGFALQSGTGAAPRLFTIHLIDANTDNP 767
 QY 687 KAHTCFNRIDIPRESYKELVKKLTAVETGFAVE 723
 DB 768 KAHTCFNRIDIPRESYKELVKKLTAVETGFAVE 804
 RESULT 7
 ID AAU87301 standard; protein; 514 AA.
 AC AAU87301;
 XX
 DT 05-JUN-2002 (first entry)
 XX
 DE Novel central nervous system protein #211.
 XX
 KW Central nervous system; CNS; autoimmune disease; rheumatoid arthritis;
 KW hyperproliferative disorder; neoplasia; cardiovascular disorder;
 KW cardiac arrest; cerebrovascular disorder; ischemia; angiodysplasia;
 KW nervous system disorder; Alzheimer's disease; AIDS; ocular disorder;
 KW acquired immunodeficiency virus; dysphagia; gastrointestinal disorder;
 KW endocrine disorder; diabetes; cancer; leukemia; neovascularisation;
 KW respiratory disorder; renal disorder; kidney failure; blood disorder;
 KW myocardial infarction; wound healing; cell proliferation; skin aging;
 KW food additive; food preservative; gene therapy.
 KW Homo sapiens.
 XX
 PD MO200155318-A2.
 XX
 PD 02-AUG-2001.
 XX
 XX 17-JAN-2001; 2001WO-US001332.
 PR 31-JAN-2000; 2000US-0179065P.
 PR 04-FEB-2000; 2000US-0180628P.
 PR 24-FEB-2000; 2000US-0184664P.
 PR 02-MAR-2000; 2000US-0186350P.
 PR 16-MAR-2000; 2000US-0189874P.
 PR 17-MAR-2000; 2000US-0190076P.
 PR 18-APR-2000; 2000US-0198123P.
 PR 19-MAY-2000; 2000US-0205515P.
 PR 07-JUN-2000; 2000US-0209467P.
 PR 28-JUN-2000; 2000US-0218866P.
 PR 30-JUN-2000; 2000US-0215135P.
 PR 07-JUL-2000; 2000US-0216647P.
 PR 11-JUL-2000; 2000US-0216880P.
 PR 11-JUL-2000; 2000US-0217487P.
 PR 14-JUL-2000; 2000US-0217496P.
 PR 26-JUL-2000; 2000US-0218295P.
 PR 26-JUL-2000; 2000US-0220963P.
 PR 14-AUG-2000; 2000US-0220964P.
 PR 14-AUG-2000; 2000US-0224518P.
 PR 14-AUG-2000; 2000US-0224519P.
 PR 14-AUG-2000; 2000US-0225213P.
 PR 14-AUG-2000; 2000US-0225214P.
 PR 14-AUG-2000; 2000US-0225266P.

CC maintain organs before transplantation, for supporting cell culture of
CC primary tissues, to regenerate tissues and in chemotaxis. The
CC polypeptides can also be used as a food additive or preservative to
CC increase or decrease storage capabilities, fat content, lipid, protein,

Query Match 57.6%; Score 2236; DB 4; Length 514;
Best Local Similarity 80.7%; Pred. No. 1,8e-210;
Matches 414; Conservative 43; Mismatches 42; Indels 14; Gaps 4;

```

QY 218 NRPCHQSDPEEGYEORTTVOGVYFLHTGTGVTWMDPRI PRDLNSVNCDELGPPLPG 277
DB 9 SETHHTPPDPEEGEORTTVOGVYFLHTGTGVTWMDPRI PRDLNSVNCDELGPPLPG 68
QY 278 WVRSTVSGRIYFVDHNNRTTQFTDPR---LHHMNNCCQCKBSPQLPSPSGSL--- 330
DB 69 WEIRKATGRYFVVDHNNRTTQFTDPRLSANLHVNQOLKQOQOQV---SLCPD 124
QY 331 EDEELPAQYERDVLQKLVRLHSLISQOPAGHCRLEVSREELFEESYRQIMRRPDL 390
DB 125 DTECLTVPYRKDVLQKLVRLHSLISQOPAGHCRLEVSREELFEESYRQIMRRPDL 184
QY 331 KRLLVFRGREGGLDYGVARERMLYLCHENLNPYGLFOYSTNTNITWLOINPSSINPD 450
DB 135 WKRLMTKRGESGLDYGVARERMLYLSHEMLNPYGLFOYSRDLITTLQINPSSAVNPE 244
QY 451 HLSTYFHVGRIMGLAVFHGHYINGFTVPFYKOLLGKEICLSDESVDPELHKSILWILE 510
DB 245 HLSTYFHVGRIMGLAVFHGHYIDGGFTLPFYKOLLGKSTILDMELVDPDLHSLWILE 304
QY 511 NDITFVLDHTECVENNAAGRLLOHLEIKNGRNVPYTEENKKEVRLVYNNRFRGIAQF 570
DB 305 NDITFVLDHTECVENNAAGRLLOHLEIKNGKSIPEVNEENKKEVRLVYNNRFRGIAQF 364
QY 571 LALQGFNELIPHLLKFPDQKELELIGLGDKIDLDWMSKNTLKGCVASDNIYVRFWQ 630
DB 365 LALQGFNEVLPQHLLKTFDEKELELIGLGDKIDVDKNTLKGCTDSNIVKAFWK 424
QY 631 AVEFDEERRARLLQFTGSTRVPLQGFKALQSTGAAPRLFTIHLIDANTDNLPRAT 690
DB 425 AVEFDEERRARLLQFTGSTRVPLQGFKALQ---GAAGRLFTIHOIDA CTNNLPRAT 481
QY 691 CFNRIDIPYSEYKLEKLTAVEETCGFAVE 723
DB 482 CFNRIDIPYSEYKLEKLTAVEETCGFAVE 514

RESULT 8
ABB61120
ID ABB61120 standard; protein, 1035 AA.
XX AC
XX ABB61120;
XX
XX 26-MAR-2002 (first entry)
XX
XX Drosophila melanogaster polypeptide SEQ ID NO 10152.
XX
XX Drosophila; developmental biology; cell signalling; insecticide;
XX pharmaceutical.
XX
XX Drosophila melanogaster.
XX
XX WO200171042-A2.
XX
XX 27-SEP-2001.
XX
XX 23-MAR-2001, 2001WO-US009231.
XX
XX 23-MAR-2000, 2000US-0191637P.
XX
XX 11-JUL-2000, 2000US-00614150.
XX
XX (PEKE ) PE CORP NY.
XX
XX Venter JC, Adams M, Li FWD, Myers EW,

```

XX WPI, 2001-656860/75.
XX N-PSDB; ABL05223.

PT New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signaling and cell-cell
PT interactions.

PS Disclosure; SEQ ID NO 10152; 21bp + Sequence Listing; English.

XX The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signaling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL01840-ABL16175), expressed DNA
CC sequences (ABL01840-ABL16175) and the encoded proteins (ABBS7737-
CC ABB78072). The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at http://wipo.int/pub/published_pct_sequences

XX Sequence 1035 AA;

Query Match 53.3%; Score 2069; DB 4; Length 1035;
Best Local Similarity 41.2%; Pred. No. 1.7e-193;
Matches 435; Conservative 104; Mismatches 145; Indels 372; Gaps 15;

```

QY 6 KTRLTYLCAKNTLAKDPRLPPFPFAKIYDGGCHSDTYKNTIDPKMNOHYLYGKT 65
DB 14 KVRITLLCAKNTLAKDPRLPPFPFAKIYDGGCHSDTYKNTIDPKMNAHYLYFLIG 73
QY 66 DSITTSVNNHKKHKKQAGFLGCVRLSNASISRLKDGORYOLDCKNPADTAVRQOI 125
DB 74 DAITTVNNQRIKH--GSGFLGCRIRAPFNIGSKAGFRLDGLKSLPDDDELVRQOI 131
QY 126 VVSLQTRDRIGTG-----GSVYDCRGLENE-----GVYV----- 155
DB 132 IISLSKQFSSGNPLAIVGSGDVRGSPEDSDSESLPGEWBRRTDNGRYYVNHATK 191
QY 156 ----- 155
DB 192 STQWRPAPQGVGSSHAATSPOQRNNTNNGSGDROAPAGTRSTCTNLMNNGHRSRL 251
QY 156 -----EDSGGRPLSCFMEEPARYDTSTGAAGAGCREVE----- 191
DB 252 SVTASDERRHSTELISVGKENTSPPTVSA-TTPGKKTSSNSSAGG--RTLEGKPT 308
QY 192 -----SPSQDQRLQAQRLANPDVRSIQTPQRPHGQSP----- 227
DB 309 NEPATPTSTTSASVRLHSND--NHVKTPKQTNHAPBSPTPTGQQYVNGNAGNG 366
QY 228 ----- 227
DB 367 STSGNGSGQAAPQASNGWTQEDAAATTSPSTTSPRRHSQSPPTNISPPASVTSAN 426
QY 228 ----- 227
DB 427 GNVSFNANSTPAGSGGSRSYTAATPGQSRSSROQGEBSSTRRRSGRTNGGNG 486
QY 228 -----LPEGEORTTVOGVYFLHTGTGVTWMDPRI 259
DB 487 GGGGGGSGQRYASAAIAAANGARFIDLPPGYEMRTTQOQGVYFYHLPVGVSTWMDPRI 546
QY 260 PRDLNS--VNCDELGPPLPGWEVRSYSGRIYFVDHNNRTTQFTDPRLHHIM----- 309
DB 547 PRDPTQHLTLDAIGPLPSGWEQKRTASGRYFVDHNNRTTQFTDPRLSGSLIQMIRGT 606
QY 310 ---NHQOQLKEPSPQLPSPS-----EG 328
DB 607 VPPTSANAGTPAPPSATPAPSAANAAPPQATPASATPTTLTTTNPPHRIYVDLPQG 666
QY 329 SLEDEELPAQYERDVLQKLVRLHSLISQOPAGHCRLEVSREELFEESYRQIMRRP 368

```



```

Db      667 LLEGADL-LPKYARDLVGKRLALRTELQTMQPSGCHRLVSNRIFEESYRLIMKRAK 725
Qy      389 DLKKRLMKRGGEGLDYGVAAEMLYLICHEMLNRYGIFQYSTNTNIMQLNPPSSIN 448
Db      726 DMKRLMKRGGEGLDYGVAAEMLYLICHEMLNRYGIFQYSTNTNIMQLNPPSSIN 785
Qy      449 PDHLSYFHFVGRIMGLAVFHGYINGGFTVPEFYKOLLKPIQLSDLESYDEPLHKSILWMI 508
Db      786 PDHLSYFHF-----LLNKRTITGDEGVDPDLHSLTMM 819
Qy      509 LENDIPVLDHTFCVEHNAFGLLOHELKNGNNAVTEENKKEYRLVYNNRMEGCIA 568
Db      820 LBNISIGIISTSVENNSFGALVHEHLKPGASIPTEENKKEYVLYVYNNRMEGCIA 879
Qy      569 QPILALOKGPELILPOLHKLKPFDOKELELIGLGDKIDLNDMKSNTLKKCVADNSIVRMF 628
Db      880 QPILALOKGPELILPOLHKLKPFDOKELELIGLGDKIDLNDMKSNTLKKCVADNSIVRMF 939
Qy      629 WQAVEFDEERRARLLQFTVGTSTVPLQGFKALQSGTGAAPRLFTIHL-IDANTDNLDPK 687
Db      940 WQAVESYSSMRARLLQFTVGTSTVPLQGFKALQSGTGAAPRLFTIHL-IDANTDNLDPK 999
Qy      688 AHTCFNRIIDIPYASYEKLYEKLTAVEETCGFAVE 723
Db      1000 AHTCFNRIIDIPYASYEKLYEKLTAVEETCGFAVE 1035

RESULT 9
AAM78877
ID      AAM78877 standard; protein; 335 AA.
XX
AC      AAM78877;
XX
XX      06-NOV-2001 (first entry)
XX
DE      Human protein SEQ ID NO 1539.
XX
XX      Human, cytokine; cell proliferation; cell differentiation; gene therapy;
XX      vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
XX      tissue growth factor; immunomodulatory; cancer; leukaemia;
XX      nervous system disorder; arthritis; inflammation.
XX
OS      Homo sapiens.
XX
XX      W02001:57190-A2.
XX
XX      09-AUG-2001.
XX
XX      05-FEB-2001; 2001WO-US004098.
XX
XX      03-FEB-2000; 2000US-00496814.
XX      27-APR-2000; 2000US-0050875.
XX      20-JUN-2000; 2000US-00598075.
XX      19-JUL-2000; 2000US-00620325.
XX      01-SEP-2000; 2000US-00654936.
XX      15-SEP-2000; 2000US-00663561.
XX      20-OCT-2000; 2000US-00693325.
XX      30-NOV-2000; 2000US-00728422.
XX
XX      (HYSE-) HYSEQ INC.
XX
XX      Tang YT, Liu C, Dramanac RT, Asundi V, Zhou P, Xu C, Cao Y;
XX      Ma Y, Zhao QA, Wang D, Wang J, Zhang J, Ren F, Chen R, Wang ZW,
XX      Xue AJ, Yang Y, Wejthman T, Goodrich R;
XX      WPI; 2001-476283/51.
XX      N-PSDB; AAK52010.
XX
XX      Nucleic acids encoding polypeptides with cytokine-like activities, useful
XX      in diagnosis and gene therapy.
XX
XX      Claim 20; Page 3847; 6221pp; English.
XX

```

```

CC      The invention relates to polynucleotides (AAK51456-AAK53435) and the
CC      encoded polypeptides (AAM73323-AAK80302) that exhibit activity elating to
CC      cytokine, cell proliferation or cell differentiation or which may induce
CC      production of other cytokines in other cell populations. The
CC      polynucleotides and polypeptides are useful in gene therapy, vaccines or
CC      peptide therapy. The polypeptides have various cytokine-like activities,
CC      e.g. stem cell growth factor activity, haematopoiesis regulating
CC      activity, tissue growth factor activity, immunomodulatory activity and
CC      activin/inhibin activity and may be useful in the diagnosis and/or
CC      treatment of cancer, leukaemia, nervous system disorders, arthritis and
CC      inflammation. Note: Records for SEQ ID NO 2110 (AAK52581), 2111
CC      (AAK52582) and 3666 (AAM80020) are omitted as the relevant pages from the
CC      sequence listing were missing at the time of publication
XX
SQ      Sequence 335 AA;
XX
Query Match      41.6%; Score 1614.5; DB 4; Length 335;
Best Local Similarity 86.2%; Pred. No. 1.5e-149;
Matches 306; Conservative 4; Mismatches 20; Indels 25; Gaps 2;
XX
Qy      168 MEPPAPYDSTGAAAGGNCRFVESPSQDQRLQARLNPVRSGLQTPQNRPHGHSPE 227
Db      1 MEPPAPYDSTGAAAGGNCRFVESPSQDQRLQARLNPVRSGLQTPQNRPHGHSPE 60
Qy      228 LPEGYEORTVQGVYFHTQGVSTWHPRIIPRLNSVNCDELGRLPFGWEVSTVSGR 287
Db      61 LPEGYEORTVQGVYFHTQGVSTWHPRIIPRLNSVNCDELGRLPFGWEVSTVSGR 120
Qy      288 IYFVDHNRRTQFTDRLHINNHQCLKEPSQPLPESGSLDEDELPQRYERDLYOK 347
Db      121 IYFVDHNRRTQFTDRLHINNHQCLKEPSQPLPESGSLDEDELPQRYERDLYOK 180
Qy      348 LKVLRLHSLQGPQAGHCRIEVSREIPEESYRQIMKRPYDLKRLNWKRRGGGLDYG 407
Db      181 LKVLRLHSLQGPQAGHCRIEVSREIPEESYRQIMKRPYDLKRLNWKRRGGGLDYG 240
Qy      408 GVAAREMVLICHEMLNRYGIFQYSTNTNIMQLNPPSSINPDHLSYFHFVGRIMGLAVF 467
Db      241 GVAAREMVLICHEMLNRYGIFQYSTNTNIMQLNPPSSINPDHLSYFHFVGRIMGLAVF 298
Qy      468 HGHYINGGFTVPEFYKOLLKPIQLSDLESYDEPLHKSILWILENDITVLDHTFC 522
Db      299 -----CVPWTLHGRC-----LHSALLQAAAGRAHPLRSGLC 330

RESULT 10
AAM13385
ID      AAM13385 standard; protein; 766 AA.
XX
XX      AAM13385;
XX
XX      10-UTL-1997 (first entry)
XX
XX      Human protein ubiquitin ligase pub2.
XX
XX      Protein ubiquitin ligase; pub2; cdc25 phosphatase; CDK kinase; p53;
XX      cell cycle; transgenic animal.
XX
XX      Homo sapiens.
XX
XX      W09712962-A1.
XX
XX      10-APR-1997.
XX
XX      04-OCT-1996; 96WO-US015930.
XX      04-OCT-1995; 95US-00539205.
XX
XX      (COLD-) COLD SPRING HARBOR LAB.
XX      Beach D, Caligiuri M, Nefeky B;
XX      WPI; 1997-226206/20.
XX

```

DR N-PSDB; AAT47041.
 XX Human and Saccharomyces pombe protein ubiquitin ligase (s) - involved in
 PT cdc35 phosphatase and p53 ubiquitination, and regulate cell growth and
 PT proliferation.
 XX
 XX Claim 1; Page 80-84, 108pp; English.
 XX
 CC Human protein ubiquitin ligases pub1 (AAW13384), pub2 (AAW13385) and pub3
 CC (AAW13386) are homologues of fission yeast pub1 (AAW13387) and were
 CC identified from cDNA clones (AAT47040-42) obtd. e.g. from a keratinocyte
 CC cDNA library. Pub polypeptides can be produced in transfected host cells.
 CC They can control the steady state level of cdc35 phosphatase, the degree
 CC of CDK kinase (e.g. cdc2) dephosphorylation and the steady state level of
 CC p53 (controlling the degree of cell cycle regulation of p53). They can
 CC regulate cell or tissue differentiation, or cell growth or proliferation
 CC by affecting other proteins, can be a specific (ant)agonist of wild-type
 CC protein function and may be used as immunogens to elicit a specific
 CC immune residue
 CC
 SQ Sequence 766 AA;
 Query Match 37.6%; Score 1461; DB 2; Length 766;
 Best Local Similarity 40.8%; Pred. No. 8,2e-134;
 Matches 326; Conservative 118; Mismatches 236; Indels 120; Gaps 20;
 QY 4 SIKITLTLCAKNAKPFPRLPDPFAKIVVDGSGCHSTVTKTLDPKNGHYDLVYG 63
 DB 7 SRRIRVILVADGILKRVDFRPPDFAVLTVDGR-QTHITTAIKKTLNPNYNETFEVAVT 65
 QY 64 KTDSTITISVMNHKRIHKQAGFLGCVALLSNALSRLKDTGYORL-DLCTLNSDPTDAV 121
 DB 66 DNSTALIVFDQKFK-KKKGGGFLGVINLRVGDVLDLAGDEMLICDLKLN-EMTVV 122
 QY 122 RQGIIVSLQ-----TRDRIGTGSGVYDCRGLLENQTYED 157
 DB 123 HKKIITINLSTALQTLQVPSAASGARTORTSINTDPOSSKSSVSRRNPASRAGSPTRD 182
 QY 158 SGPQ-----RPLSCFMEB---PAPY---TGSTG-----AAA 182
 DB 183 NAPAASPASSERITSSFEQYGRLPPEMERCTDNLGTYVDNNTSTWIRNLSVA 242
 QY 183 GGCNCRFVESPS-----QDRLQALRNPDVRSGLQTPQNRPHGOSPELPEGYEQ 234
 DB 243 GAAAEHLSSASNAVTEGVQPSSSNARTEASVLTSNATTAG-----SGELLPMQEQ 296
 QY 225 RTTVGGGVYELHTQGVSTWMDPRIRPDLNSV-----CDELGPPLPGMEVYST 283
 DB 297 RTTPGRPFVDMNRTTITWDPDRROQTIRSYGCPNNTATIQQPVSGPLPSGEMERLT 356
 QY 284 VSGRIYFVNNHNTQFTDPRLHIMNHQCLKEPSQPLPSPSGSEDEELPAQRYERD 343
 DB 357 NTARVYFVDHNTKTTMDPR-----LPS--SL-DQVVP--QYKRD 392
 QY 344 LVQKIKVLRHELSLQOPQAGHCRIEVREIPEBSYRQIMQMRKDLKRLMVFRRBEG 403
 DB 393 FRRKIYVLSQPAL-HPLPGQCHIKVRNHLFEOSYALIMQASATDLKRLMIFRDEBDG 451
 QY 404 LDYGVAREWLYLCHENLNPYYGLFOYSTDNIVYLQINPSSINPDLSYFHVGRING 463
 DB 452 LDYGLSEYFPLLSHEMNPFCYCFEYSSVDNTLQINPNSGINPELNAFKRIGVIG 511
 QY 464 LAVFHGYINGFTVFFYKOLGKFIQSLDESVDPELHKLVLWLENDITPVIDHTFCV 523
 DB 512 LAIFRRFVDAFVVSFYKMLQCKVTLCQDMSMDAEYRSLVWILNDITGVLDLTFVS 571
 QY 524 ENNAAGRILOHLEKNGSNVPEENKKEYRLVYNNRPFMGIGAPFLALQKGNELIPQ 583
 DB 572 EDNCGEAVYITDKRNGNRIETENKREYDVLTVW-IQRIEQRPAHFEQSEILIPQ 630
 QY 584 HLKFPDQKLELLIIGADKIDLNDWKSNTLKAICVADSNIVRPFQAVETFEREARL 643
 DB 631 ELINVFDERLELLIGGISLIDMEDWKKGXDYSSENDQITIKPFWELMDWMSRKSRL 690

QY 644 LQFVGTSTVPLQGFYKALQSGTGAGPRFLTILIDANTDLPKATGCFNRIDIPPEBSY 703
 DB 691 LQFTTSTSLPNNGFDDLOGSD---GPRKFTLEKA-GERNKLPKXATGTCNRDLPLPYNSX 746
 QY 704 EKLYEKULTRVETGCFAYE 723
 DB 747 KDLDRKLSIAVEETIGFGQE 766
 RESULT 11
 ID ABP73459
 ID ABP73459 standard; protein; 832 AA.
 XX
 AC ABP73459,
 DT 30-JAN-2003 (first entry)
 XX
 DE Candida albicans essential protein SEQ ID NO 7296.
 XX
 XM Fungus; yeast; tetracyclin; promoter; GRACE strain; biosynthesis;
 XM signal transduction; DNA replication; cell division; growth;
 XM proliferation; Candida albicans; fungicide; antifungal.
 XX
 OS Candida albicans.
 XX
 PN M0200253728-A2.
 XX
 PD 11-JUL-2002.
 XX
 PF 26-DEC-2001, 2001WO-US049486.
 XX
 PR 28-FEB-2000, 2000US-0259128P.
 PR 20-FEB-2001, 2001US-0079202P.
 PR 22-AUG-2001, 2001US-0314050P.
 XX
 PA (ELITR) ELITRA PHARM INC.
 XX
 PI Roemer T, Jiang B, Boone C, Bussey H, Ohlsen KL;
 XX
 DR WPI, 2002-566694/60.
 DR N-PSDB; AB232009.
 XX
 PT Constructing strains for identifying gene products as effective targets
 PT for therapeutic intervention, by inactivating in the strain one allele of
 PT a gene and placing other allele of the gene under conditional expression.
 XX
 PS Claim 44; SEQ ID NO 7296; 167bp + Sequence Listing; English.
 XX
 CC The invention relates to constructing (M1) a strain of diploid fungal
 CC cells in which both alleles of a gene are modified, comprising modifying
 CC one allele by insertion or replacement by a cassette having an
 CC expressible selectable marker and modifying other allele by
 CC recombination of a promoter replacement fragment with a heterologous
 CC promoter, so that expression of the second allele is regulated by the
 CC promoter. (M1) is useful for constructing a strain of diploid fungal
 CC cells having both alleles modified are useful for identifying a gene that
 CC is essential to the survival or growth of a fungus, a gene that
 CC contributes to the virulence and/or pathogenicity of a fungus, a gene
 CC that contributes to the resistance of a diploid fungus to an antifungal
 CC agent, an antifungal agent that inhibits the growth of a diploid fungus
 CC and for identifying a therapeutic agent for treatment of a mammalian
 CC disease. (M1) is useful for identifying a compound which modulates the
 CC activity of a gene product, preferably enzymatic activity, carbon
 CC compound catabolism, biosynthetic, transporter, transcriptional,
 CC translational, signal transduction, DNA replication and cell division
 CC activity. The method is useful for identifying a compound having the
 CC ability to inhibit growth or proliferation of C. albicans cells and for
 CC treating infection by C. albicans. The present sequence is that of an
 CC essential Candida albicans protein used in the method of the invention.
 CC Note: The sequence data for this patent is not represented in the printed
 CC specification but is based on sequence information supplied to Derwent by


```

Db 82 LNPYNEMFDRVNEDESLAIQIFDQKF-KKQDQGLGVINVAIGVIDLQMGDGESL 140
QY 94 ----SNAISRLK-----DTGYQLDLCKLNPSDTAVRGQIVVSLQ----- 131
Db 141 PIRHSCDVSRLTFRLLLETEMLTDLKSN--DNLVHVGKLIINLSNLSNTPNTQANG 198
QY 132 --RDRIGTGGVDCRGLLEN-----EGTVIED--SGPG-----RPLSC 166
Db 199 LHRSHVOSSTS-----SGLVPQVAPSSGHPAASGTAPVDPASPSLNPQRPVSTTRFSS 254
QY 167 FMEBPA-----PYTDSGAAAG--GNCRFVE-----SP 193
Db 255 AAPASAGAASVNSHGSRNLSFEDSQRLPAGWEREDNLGRTYVDHNTTITWTRP 314
QY 194 SODQLQRLRN-----PDVR--GSLQTPQNRPHGOSP----- 226
Db 315 SSNTYNEHAQRSQREANWOLERRAHSQRLPEDRTGANSPLPSSQQAHTPPAGGSANAV 374
QY 227 -----ELPEGYEORTTVQGVFLHTQTGVTWHDPR-----IPRLNSV 266
Db 375 SMWATGATTAGTGLPQWEGORTPEGRPFYVDHNTTITWDPFRQOYRMGQANGT 434
QY 287 N-----CDELGLPQWGVRSVSGRIYFVDHNNRTTQFTDRLHHIMHQCQLKEPS 319
Db 435 NTTIQQPVSLPGLPSCWENLNTARVYFVDHNTTITWDDPR----- 479
QY 320 QPLPLPSEGSLEDELPQAQYERDLVKLVRLHLSLOQPOACHIEVSRSEIPEESY 379
Db 480 ----LPS--SL-DQGV--QYKDRFRKLIYFRSQPAL-RIMSGQCHVKVRRNIFEDSY 529
QY 380 ROIMKMRPKDLKLMKFRCEGLDGGVAREWLYLLCHEMLNPFYGLFOYGTDMLYL 439
Db 530 AEIMQASADLKKLMKFDEGDLGYGLSGEFFLLSHEMFNPFCLFEYSAHDNYTL 589
QY 440 QINPDSSINPHLSYFHFVGRIMGLAVFHGYINGGTVPFYKQLLQKPIQLSLSVDP 499
Db 590 QINPHSGVNPHELYFPIGRVWGLAIFRRFLDSFFIGAFYKMLKVKVLSQDMQGVDE 649
QY 500 ELKSLVNLBNDITPVLDTFFVCHNAFGRILQHELKPGRNVPVTEENKKEVVRLYVN 559
Db 650 DLHRLNTWNDIEGVLELTFSDDEKFGERTIDLPKGRDIPVTNEKAYFVRLVTE 709
QY 560 WRFMEGIEAQLALQGFNELIPQHLKLPDQKELEIIGLKDIDNDKSNTRLKHCV 619
Db 710 WKIVKRVEEQFNAFWSGFNELIPADLVNVDERELELLIGIADIDVDDKKHTDYRGY 769
QY 620 ADSNIVRFWQAVETFDERRARLLQVGTGSTRVPLQGFALOGSTGAAGPRLFTIHLID 679
Db 770 ESDEVIONFWKIVRSWDAEQKSRLLQFTTGTSRIPVNGFKDLOGSD--GPRRFTIEK-S 825
QY 680 ANTDLNPKAHTCFNRIDIPPYSEYKELYELLTAVEETCGFAVE 723
Db 826 GDPAALPKSHGTGNRLDLPYKSYETLEHKVNSIAVEETLGFQGE 869

```

RESULT 13

AAY30949

ID AAY30949

XX AAY30949 standard; protein; 854 AA.

XX AAY30949;

XX AAY30949;

XX AAY30949;

XX AAY30949;

XX AAY30949;

XX AAY30949;

XX AAY30949;

XX AAY30949;

XX AAY30949;

XX AAY30949;

XX AAY30949;

XX AAY30949;

XX AAY30949;

XX AAY30949;

XX AAY30949;

XX AAY30949;

XX AAY30949;

XX AAY30949;

```

XX Mus sp.
OS WO9940201-A1.
XX 12-AUG-1999.
XX 02-FEB-1999; 99WO-GB000353.
XX 05-FEB-1998; 98US-0073839P.
XX 30-APR-1998; 98US-00070060.
XX (ZENE ) ZENECA LTD.
XX Hustad CM, Ghildyal N;
XX WPI; 1999-508506/42.
XX New human proteolytic accessory enzyme and its modulators useful for
XX treating disease conditions like inflammation or autoimmune diseases.
XX Disclosure; Fig 4; 95pp; English.
XX This invention describes a novel human polynucleotide (I) which encodes a
XX E3 ubiquitin protein ligase, h-E3 UPL (II). The products of the invention
XX have antiinflammatory, immunosuppressive, neuroprotective, cytostatic,
XX antiarthritic, immunomodulator, antidiabetic, antiviral and cardiac
XX activity. The products of the invention can also be used for treating
XX patients with disorders mediated by the biological and/or pharmacological
XX activity of h-E3 UPL. The nucleic acid sequences encoding E3 UPL are used
XX in expression systems as assay for agonists and antagonists for the E3
XX UPL protein. The E3 UPL protein is used in screening assays to identify
XX blockers and antagonists. They are also used in gene therapy. Specific
XX modulation of biological and/or pharmacological activity of novel h-E3
XX UPL via administration of a modulator or heterologous expression, is used
XX for treating physiological conditions like inflammation, autoimmune
XX diseases, neurological disease, apoptosis, endothelial cell physiology
XX (e.g., proliferation, differentiation), peripheral vascular disease,
XX angiogenesis, cancer, hematopoietic disorders, arthritis, cachexia,
XX leukemia, pulmonary disorders, diabetes and viral infection. The
XX nucleotide sequences which encode h-E3 UPL may also be employed in
XX analysis to map chromosomal location e.g., screening for functional
XX association with disease markers. They are also used as screening tools
XX in the identification of appropriate human subjects and patients for
XX therapeutic clinical trials. The sequences can also be used to detect the
XX presence of the mRNA transcripts in a patient or to monitor the
XX modulation of transcripts during treatment. This sequence represents the
XX mouse E3 ubiquitin protein ligase protein described in the method of the
XX invention
XX Sequence 854 AA;

```

```

Query Match 33.9%; Score 1318; DB 2; Length 854;
Best Local Similarity 35.4%; Pred. No. 1.2e-119;
Matches 316; Conservative 127; Mismatches 229; Indels 220; Gaps 29;

```

```

QY 6 KIRLVLCAL-NLAKODFFRLPDPFAKIVVDGSGQCHSTDTVKNTLDPKXNHVDLYGK 64
Db 9 QLIQTVISAKLKENKQWFG-PSFYVEVTVD--GQSKKTEKCNNTNSPKWQPIVTVTP 65
QY 65 TDSITISVNHKKHKQAGFLGCVRL-----LSNAISRLKDTGYOR----- 107
Db 56 TSKLCFRVWSHQTL---KSDVLLGTAGLDIYETLKNMKLEEVVMTLQLVGDKEPTETM 122
QY 108 --LDICKLNPSDPAVRGQIVVS-----LQTRD--RIGTGS-----VDCR 145
Db 123 GDLVYC---LDGLQVEAEVVTNGETSCSESTQNDGCRTRDDTRVSTNGSEDFEVAAS 178
QY 146 G-----LLENEGTVVEDSGRPLSCFVEEAPYV-----DSTG 179
Db 179 GUNKANGNPSLSNGG--FKPSRPPRP-----SRPPPTPRPASVNGSPSINSUSDG 231
QY 180 AA-----AGGNCRFVESPQ-----DORLOAQR 203

```

Murine E3 ubiquitin protein ligase protein.

E3 ubiquitin protein ligase, h-E3 UPL; antiinflammatory; antidiabetic; immunosuppressive; neuroprotective; cytostatic; antiarthritic; cardiac; immunomodulator; antiviral; treatment; screening; gene therapy; cancer; inflammation; autoimmune disease; neurological disease; apoptosis; endothelial cell; proliferation; differentiation; angiogenesis; cachexia; peripheral vascular disease; hematopoietic disorder; arthritis; leukemia; pulmonary disorder; diabetes; viral infection; murine.

```
Db 232 SSTGSLPTNTNTSTSEATSGLIPTISGGGPRELNTVSOAPLPGEORV----- 287
Qy 204 LRNPVVRGSL-----OTPNRPHGHQSPELPEGEORITVQGOVYFLHTQTGVSTWH 255
Db 288 -----DOHGRVYVDHVEKETTWDPR-----PLPPGWERVDNMGGIYVDFHTTTWQ 338
Qy 256 DPRIP-----RDLE-----NSVNCDELGPLPGEWVR 281
Db 339 RPTLESVRNVEQWLORLOQAMQONRFYVGNQDLFATSKQNKFEPLPLPGEWKR 398
Qy 282 STVSGRIYFVDHNRRTQTDRLHMHMHQOLKEPSOPLPL----- 324
Db 399 TDSNGRVIYFNHNRITQWEDR-----SQGLNE--KPLPEGWERTVDGIFYVDH 450
Qy 325 -----PSEGSLEDELPQRYERDLVQKLKVLK-----HELSQLQOQAGHCRIVERSE 372
Db 451 NRRATTYIDPRTGKSALDNGPQIAYVRDFKAKVQYFRWCQQLAMPQ-----HIKITVTRK 506
Qy 373 EFPEESYRQIMKMRPKDLKRLMKVRGREGIDYGGVAREWYLLCHEMLNYYGLFQYS 432
Db 507 TUFEDSFQOIMSPQDLRRRLWVIFPGBEGLDYGGVAREWFLLSHEVINMYCLFEYA 566
Qy 433 TDNIYMLQINPDSISNPDLHSYFHFVGRIMGLAVFHHYINGFTVPFYKQLLGRPIQLS 492
Db 567 GKDNVCLQINPASYINPDHLKYFRFTRGFIAALFHGKFDITGFSLPFYKRILNKPVGLK 626
Qy 493 DLESVDPELHKSILWLENDITPV-LDHTFCVEHNAFGRILHCLKPNGRNVPVTEENKK 551
Db 627 DLESIDPEFYNLSIWKENNIBCEGLEMVFSVDKEILGSIKSHDLKPNNGNTLVTEENKE 686
Qy 552 EYVRLVYVNRVGRIGBAQFLALQKGNELIPOHLLKPFDPKLELILGLDKIDLNDKKS 611
Db 687 EYIRVMAEVLRSVGEVQEQAFEGNEILPQOYLQYFDAKEVLELLCQWQIDLNDWR 746
Qy 612 NTRLKHCVADSNIVRFWQAVETFDSEERARLLQFVTGSTRVPLQGFKALQSGTGAAGR 671
Db 747 HAIYRHYRTSKQIMWFQVKEIDNEKMRLLQFVTGTCRLPVGGFADLNGSN--GPQ 803
Qy 672 LFTIHLIDANTNLKPAKCFNRIDIPYVESYEKLEKLLTAAVEECGFAVE 723
Db 804 KPCIEKV-GKENWLPKSHTCFNRLDLPYKSYBQLAKELLFALTEEGFGQE 854

RESULT 14
AAV30948
ID AAV30948 standard; protein; 852 AA.
XX
AC AAV30948;
XX
DT 21-OCT-1999 (first entry)
XX
DE Human E3 ubiquitin protein ligase protein.
XX
KW E3 ubiquitin protein ligase; h-E3 UPL; antiinflammatory; antidiabetic;
KW immunosuppressive; neuroprotective; cytostatic; antiarthritic; cardiac;
KW immunomodulator; antiviral; treatment; screening; gene therapy; cancer;
KW inflammation; autoimmune disease; neurological disease; apoptosis;
KW endothelial cell; proliferation; differentiation; angiogenesis; cachexia;
KW peripheral vascular disease; hematopoietic disorder; arthritis; leukemia;
KW pulmonary disorder; diabetes; viral infection; human.
XX
OS Homo sapiens.
XX
Key Location/Qualifiers
Domain 275..306
/label= WWI
/notes= "WW protein interaction domain I"
FT
FT 307..340
/label= WWII
/notes= "WW protein interaction domain II"
FT
FT 386..420
/label= WWIII
```

```
FT Domain /note= "WW protein interaction domain III"
FT 427..460
FT /label= WWIV
XX /note= "WW protein interaction domain IV"
XX WO3940201-A1.
XX
PD 12-AUG-1999.
XX
PF 02-FEB-1999; 99WO-GB000353.
XX
PR 05-FEB-1998; 98US-0073839P.
XX 30-APR-1998; 98US-00070060.
XX (ZENE ) ZENECA LTD.
XX
XX Hustad CM, Ghildyal N;
XX
XX WPI; 1999-508506/42.
XX N-PSDB; AAZ09235, AAY09236.
XX
XX New human proteolytic accessory enzyme and its modulators useful for
XX treating disease conditions like inflammation or autoimmune diseases.
XX
XX Claim 1; Fig 3; 95pp; English.
XX
CC This invention describes a novel human polynucleotide (I) which encodes a
CC E3 ubiquitin protein ligase, h-E3 UPL (II). The products of the invention
CC have antiinflammatory, immunosuppressive, neuroprotective, cytostatic,
CC antiarthritic, immunomodulator, antidiabetic, antiviral and cardiac
CC activity. The products of the invention can also be used for treating
CC patients with disorders mediated by the biological and/or pharmacological
CC activity of h-E3 UPL. The nucleic acid sequences encoding E3 UPL are used
CC in expression systems as assay for agonists and antagonists for the E3
CC UPL protein. The E3 UPL protein is used in screening assays to identify
CC blockers and antagonists. They are also used in gene therapy. Specific
CC modulation of biological and/or pharmacological activity of novel h-E3
CC UPL via administration of a modulator or heterologous expression, is used
CC for treating physiological conditions like inflammation, autoimmune
CC diseases, neurological disease, apoptosis, endothelial cell physiology
CC (e.g., proliferation, differentiation), peripheral vascular disease,
CC angiogenesis, cancer, hematopoietic disorders, arthritis, cachexia,
CC leukemia, pulmonary disorders, diabetes and viral infection. The
CC nucleotide sequences which encode h-E3 UPL may also be employed in
CC analysis to map chromosomal location e.g., screening for functional
CC association with disease markers. They are also used as screening tools
CC in the identification of appropriate human subjects and patients for
CC therapeutic clinical trials. The sequences can also be used to detect the
CC presence of the mRNA transcripts in a patient or to monitor the
CC modulation of transcripts during treatment. This sequence represents the
CC human E3 ubiquitin protein ligase protein described in the method of the
CC invention
XX
XX Sequence 852 AA;
XX
Query Match 33.5%; Score 1301; DB 2; Length 852;
Beat Local Similarity 34.7%; Pred. No. 5,7e-118;
Matches 308; Conservative 126; Mismatches 241; Indels 212; Gaps 26;
Qy 6 KIRUTVLCAG-NLAKGPPFRUPDPFAKIVVDGSGQCHSTDTKNTLDPKKNQHYDLVYVGK 64
Db 9 QLOITVISAKLENKQWFG-PSPIVEVVD--QQSKTEKCNNTNFPKWQPLTIVTVP 65
Qy 65 TDSITISVNNHKTHKK--QGAGFLGCVRLLSNAISRLKDT-----GYQR-----L 108
Db 66 VSKLHFRVMSHQTLKSDVLLGTAALDIYETLKSNNMKLEEVVVTLQLGDKREPETIGDL 125
Qy 109 DLCKLNPSDITDAVRGQIVVSLQT-----RDRIGTGS-----VDDC 144
Db 126 SIC-----LDGLQLESEVVNTGETTCSESASQNDGSKRSKQETRVSTNGSDDDPEPAGAG 181
Qy 145 RGL-----LENEGTVYEDSGPGRP-SCFMEEPAPYTDSTGAAGGNCNRFVES----- 192
```

Db 182 RRVSGNNSFSLNNGG--FKPSRPRP-----SRPPPTPRPAPVNGSPATSSDSGSST 234
 QY 193 -----PSQDRLQAORLNPDP 208
 Db 235 GSLPPTNTNTSEGATGLIPLTISGSGRPLNPVTQAPLPGWEQVRV-----D 286
 QY 209 VRGSL-----QTPQNRPHGHSPELPEGYEORTTVQGVYFLHTQTGVSTWHDPRIP 260
 Db 287 QHGRVYVVDHVEKRTWDRPE-----PLPFGWERRVDNNGRIYVVDHFTTRTTWQRTTLE 341
 QY 261 -----RDLE-----NSVNCDELGLPLPGWEVRSTVSG 286
 Db 342 SVRNYEQWLQSQLOGAMQFNQRFYIGNODLFATSSQSEKDFPLGLPLPGWEKRTUSNG 401
 QY 287 RIYFVDHNNRTQFTDPLHMHMHQCOLKEPSPOLPL----- 324
 Db 402 RYFVNNHNRITQWEDPR-----SQQLNE--RPLPEGWEMRFTVDGIPYFVDHNNRTT 453
 QY 325 ----PSEGSLEDELPACQRYERDLVOKLVLR---HELSQLQOQAGHCRIEVSREEIFEE 377
 Db 454 TYIDERTKSALDNGPQIAYVRDPKAKQYFRCQQLAMPQ-----HIKITVTKILFED 509
 QY 378 SYRQIMKWRPKDKRLKRLMYKRGEGLDYGGVAREWLYLLCHEMLNPYGLFQVSTDNII 437
 Db 510 SPQIMSPSQDLRRRLRWIIPGBEGLDYGGVAREWFLLSHEVLNPNMYCLFEYAGKDN 569
 QY 438 MLOINPDSINPDHLSYFHFVGRIMGLAVRHGVTGTPVFKYKOLLGKPIQLSDLESV 497
 Db 570 CLQINPASYIIPDHLKUYFRFIRIAMALFGKFDITGFSLFFRYRILNPKVPLGLKLESI 629
 QY 498 DPDLKSLWILENDITPV-LDHTFCVEHNAFRILQHELKPNGRNVPVTEENKKEYVRL 556
 Db 630 DPEFYNSLIWKNNIBECDELMVFSVDKEILGSHDLKPNNGNINLVTEENKEEYIRM 689
 QY 557 YNWRPFMGIAQFALQKGNELIPQHLKPFQKLELILGGLDKIDLDNWSNTRLK 616
 Db 690 VAEWRLSRGVEEQTAFFEGNEILPQOYLQYFDAKLEVLVLLCGMBIDLNDWQHAIYR 749
 QY 617 HCVADSNIVRPFQAVETFDSEERARLLQVGTSTRVPLQGFKALQGSTGAAGPRLFTIH 676
 Db 750 HYARTSKQIMFWQFVKEIDNEKRMLLQVYTGTCRLPVGGFADLMGNN---GPKQFCIE 806
 QY 677 LIDANTDLNPKATCFNRIDIPVSEYKLYEKLITAVEETCGFAVE 723
 Db 807 KV-GKENWLPKRSHTCFNRLDLPYKSYEQKLEKLLFAIETEGRGQE 852

RESULT 15
 ABR64210
 ID ABR64210 standard; protein; 903 AA.
 AC
 XX
 DT 15-OCT-2003 (first entry)
 XX
 DE Angiogenesis protein BNO104.
 XX
 KW Cystostatic; antirheumatic; antiarthritic; antidiabetic; ophthalmological;
 KW antipsoriatic; antiarteriosclerotic; cardiant; vasotropic; angiogenesis;
 KW gene therapy; vasculature; cancer; rheumatoid arthritis; psoriasis;
 KW diabetic retinopathy; cardiovascular disease; atherosclerosis;
 KW ischemic limb disease; coronary artery disease.
 XX
 OS Homo sapiens.
 XX
 FN WO2003027285-A1.
 XX
 PD 03-APR-2003.
 XX
 PF 19-SEP-2002; 2002WO-AU0001282.
 XX
 PR 27-SEP-2001; 2001AU-00007973.
 PR 27-SEP-2001; 2001AU-00007974.

PR 11-OCT-2001; 2001AU-00008210.
 PR 29-OCT-2001; 2001AU-00008532.
 PR 13-NOV-2001; 2001AU-00008838.
 PR 28-AUG-2002; 2002AU-00951032.
 XX
 PA (BION-) BIONOMICS LTD.
 XX
 FI Gamble JR, Hahn CN, Vadas MA;
 XX
 DR WPI; 2003-354655/33.
 DR N-PSDB; ACF34485.
 XX
 PT New angiogenic genes and polypeptides, useful for diagnosing,
 PT prognosticating or treating an angiogenesis-related disorder, e.g.
 PT cancer, rheumatoid arthritis, diabetic retinopathy, psoriasis or
 PT cardiovascular diseases.
 XX
 PS Claim 15; SEQ ID NO 145; 90pp; English.
 XX
 CC The invention relates to the isolation of novel genes (ACF34446-ACF34559)
 CC encoding proteins (ABR64180-ABR64281) involved in the process of
 CC angiogenesis. The nucleic acid molecules are useful in identifying and/or
 CC obtaining full-length human genes involved in an angiogenic process. The
 CC nucleic acid molecule, polypeptides or complexes encoded, cells or
 CC genetically modified non-human animals derived from these are useful for
 CC the screening of candidate pharmaceutical compounds used in treating
 CC angiogenesis-related disorders. They are also useful for diagnosing,
 CC prognosticating or treating an angiogenesis-related disorder, which
 CC involves uncontrolled or enhanced angiogenesis or is a disorder in which
 CC a decreased vasculature is of benefit (e.g. cancer, rheumatoid arthritis,
 CC diabetic retinopathy, psoriasis or cardiovascular diseases such as
 CC atherosclerosis), or involves inappropriately arrested or decreased
 CC angiogenesis or is a disorder in which an expanding vasculature is of
 CC benefit (e.g. ischemic limb disease or coronary artery disease). The
 CC modulator of expression or activity of the polypeptide encoded by the
 CC nucleic acid sequence is useful for manufacturing a medicament for the
 CC treatment of an angiogenesis-related disorder. This sequence corresponds
 CC to one of the novel angiogenic protein
 XX
 SQ Sequence 903 AA;

Query Watch 33.2%; Score 1288.5; DB 6; Length 903;

Best Local Similarity 33.8%; Pred. No. 1.1e-116;
 Matches 312; Conservative 128; Mismatches 239; Indels 245; Gaps 29;

QY 6 KIRLTVLCAK-NLAKKDFRPLDPFPAKIVDGGSGCHSTDTVKNLTDPKWNQHYDLYYVK 64
 Db 19 QLQITVISAKLKNKNWFG-PSYVEVTVD--GQSKTEKCNNTNSPKWQPLTVITVP 75
 QY 65 TDSITISVNHKKIHK-----KQAGFLG-- 88
 Db 76 VSKLHFRVWSHQTLKSDVLLGTAALDIYETLKNWNKLEEVVVVLQLGDKPEPTTIGDL 135
 QY 89 --CV-----RLLSNAISRLKDTGYQ---RLDL-----CKL-----NPSDTPAV 121
 Db 136 SICLDGLQLESEVINGETTCSENGVSLCLPLFECNSAISAHCNLCPLGSDSPISASRV 195
 QY 122 RGQIVVS-----LQTRD--RIGTGS-----VVDGRGL-----LENEGTVVEDSGP 160
 Db 196 AGFTGASQNDGDSRSDKDETRVSTNGSDDPEDAGENRRVSGNNSPSSLNNG--FKPSRP 253
 QY 161 GRPLSCFMEEPAPYTDSTGAAGGNCRCFVS----- 192
 Db 254 PRP-----SRPPPTPRPASPVSNGSPSATSSDSGSTGSLPPTNTNTYSEGATGLIIP 308
 QY 193 -----PSQDRLQAORLNPDPVRGSL-----QTPQNRPHG 223
 Db 309 LATSIGSGRPLNPVTQAPLPGWEQVRV-----DQGRVYVVDHVEKRTWDRPE-- 358
 QY 224 QSPPELPEGYEORTTVQGVYFLHTQTGVSTWHDPRIP----- 260
 Db 359 ----PLPFGWERRVDNNGRIYVVDHFTTRTTWQRTTLESVRNYEQWLQSQLOGAMQFN 415

This Page Blank (uspio)

Result No.	Score	Query %		Length	DB	ID	Description
		Match	Length				
1	1492.5	38.4	767	2	T37545	ubiquitin-protein	
2	1471	37.9	766	1	S66562	ubiquitin-protein	
3	1437	37.0	815	2	T43744	probable ubiquitin	
4	1377	35.5	809	1	S43217	ubiquitin-protein	
5	1341	34.5	786	2	T35855	ubiquitin-protein	
6	1241.5	32.0	820	2	T46412	ubiquitin-protein	
7	1227.5	31.6	887	2	S70642	ubiquitin ligase N	
8	1051.5	27.1	708	2	T83196	NEDD-4 ORF - mouse	
9	1026.5	26.4	671	2	T37900	probable ubiquitin	
10	854.5	22.0	3839	2	T49799	related to TOM1 pr	
11	848	21.8	3327	2	T37964	probable ubiquitin	
12	784	20.2	4056	2	H96599	ubiquitin-protein	
13	780.5	20.1	1126	2	T01491	protein F4UJ6.10	
14	769	19.8	3368	2	S69625	ubiquitin-protein	
15	669	17.2	889	2	T0274	hypothetical protein	
16	627.5	16.2	310	2	T2646	hypothetical protein	
17	604.5	15.6	221	2	T51886	DNA binding protein	
18	530.5	13.7	874	2	A39920	hypothetical protein	
19	484	12.5	807	2	T40821	B6-associated prot	
20	481.5	12.4	892	2	S57055	probable ubiquitin	
21	471	12.1	1142	2	T46155	probable membrane	
22	467	12.0	1083	2	A38919	hypothetical protein	
23	466	12.0	873	2	H95114	hypothetical protein	
24	451	11.6	1029	2	T38951	polyubiquitin-like	
25	449.5	11.6	1054	2	B38939	probable ubiquitin	
26	444	11.4	1866	2	T26753	hypothetical protein	
27	418	10.8	1001	2	T20373	hypothetical protein	
28	400.5	10.3	655	2	T41750	probable ubiquitin	
29	382.5	9.8	4836	2	T14346	herc2 protein - mo	

Db 66 DNSTIAIQVFDOKKF-KKKGGQGLGVINLRVGVLDLALGDEMLTRDLKKN--ENTVV 122
QY 122 RGQIVVSLQ-----TRDRITGGSVVDCRGLLENGTVYED 157
Db 123 HGKIIINLSTTAQSTLQVPSAASGARTQRTSITNDPQSSQSSSVSRNPASSRAGSPTRD 182
QY 158 SGPG-----RPLSCFME-----PAPY-----TDSTG-----AAA 182
Db 183 NAPAASPASSPRFTSSFDQGLRPPGWERDNLGRYVYDHNTRSTTWIRNLSVA 242
QY 183 GGCNCRFVSPS-----QDRLQARLNPDPVRSGLQTPQNRPHGHQSPBLPGYEQ 234
Db 243 GAAAELHSSASSANVTGQVPSNAARTEASVLTSNATTAG-----SGELPPGWEQ 296
QY 235 RTTVQGVYVLTQTVSTVHDPRI PRDLNSV-----CDELGPLPGWEVRS 283
Db 297 RYTPGRIYFVDHNTRITTTVDPRQYIRSYGGPNNTIQQQVPSQGLPSPGWEMLT 356
QY 284 VSGRIYFVDHNTRITTTVDPRQYIRSYGGPNNTIQQQVPSQGLPSPGWEMLT 343
Db 357 NTARVYFVDHNTRITTTVDPRQYIRSYGGPNNTIQQQVPSQGLPSPGWEMLT 392
QY 344 LVOKLVLRHLSLQOQAGHCRIEVSREIFBESYRQIMKMPKDLKKRLMKVFRGEG 403
Db 393 FRKLIYFSLQAL-HPLPGQCHIKVRNHI FEDSYABIMRQSATDILKKRLMKVFRGEG 451
QY 404 LDYGGVAREWLYLICHEMLNPYYGLFOYSTDNIMQLQINPDSSINPDHLSYFHFVGRIMG 463
Db 452 LDYGLSREYFLLSHMFNPFYCLFEYSVDNYTLQINPHSGINPEHLNFKFGRVIG 511
QY 464 LAVFHGYINGGTVVPFKYKOLLGKPIQLSDLESVDPELHKSILWILENDITPVLDHTFCV 523
Db 512 LAIFHRFVDFAFFVYFVKMLQKVTLDQWESDAEYVESLWILNDITGVLDLTFV 571
QY 524 EHNAPGRILQHELKPNRNPVTENKEEYVRLYVNRFRMRGIEAOPFLAQKGNELIPQ 583
Db 572 EDNCFGEVITDLKPNRNIETVENKEEYVRLYVNRFRMRGIEAOPFLAQKGNELIPQ 631
QY 584 HLLKPPDQKELIIGGLDKIDLNDKSNTRKHKVADSNIVRWFQAVETFDERRARL 643
Db 632 ELINVFDERELELLIGGSEIDMEDWKKHDKYRSYENDQIKWFLMDWSENEKKSRL 691
QY 644 LQFTVGTSTVPLQGFALQGSTGAAGPRLFTIHLIDANTNLPAHTCFNRDIPPYESY 703
Db 692 LQFTGTSTVPLQGFALQGSTGAAGPRLFTIHLIDANTNLPAHTCFNRDIPPYESY 747
QY 704 EKLKLTAVETTCGFAVE 723
Db 748 KDLHKLTAIVETTCGFAVE 767
RESULT 2
S66562
ubiquitin-protein ligase (EC 6.3.2.19) - fission yeast (Schizosaccharomyces pombe)
N/Alternate names: E6-AP-like protein ubiquitin ligase
C/Species: Schizosaccharomyces pombe
C/Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 03-Jun-2002
C/Accession: S66562; T45159
R/Nefsky, B.; Beach, D.
EMBO J. 15, 1301-1312, 1996
A/Title: Publ acts as an E6-AP-like protein ubiquitin ligase in the degradation of cdc25
A/Reference number: S66562; MUID:96205868; PMID:8635463
A/Accession: S66562
A/Status: preliminary; nucleic acid sequence not shown
A/Molecule type: mRNA
A/Residues: 1-766 <NEF>
A/Cross-references: GB:U66716; NID:g1519443; PIDN:AAB07514.1; PID:g1519444
submitted to the EMBL Data Library, August 1996
A/Description: Publ acts as an E6-AP-like protein ubiquitin ligase in the degradation of
A/Reference number: 222935
A/Accession: T45159
A/Status: preliminary; translated from GB/EMBL/DBDJ

A/Molecule type: mRNA
A/Residues: 1-766 <NE2>
A/Cross-references: EMBL:Y07592; PIDN:CAA68867.1
C/Genetics:
A/Genes: publ
C/Supfamily: yeast ubiquitin-protein ligase; WW repeat homology
C/Keywords: ligase
F/205-242/Domain: WW repeat homology <WW1>
F/288-325/Domain: WW repeat homology <WW2>
F/345-382/Domain: WW repeat homology <WW3>
Query Match 37.9%; Score 1471; DB 1; Length 766;
Best Local Similarity 40.9%; Pred. No. 6.7e-99;
Matches 327; Conservative 118; Mismatches 235; Indels 120; Gaps 20;
QY 4 SIKIRLTVLCAKNAKDPFRLLPDPFAKIVDVGSGQCHSDTNTKNTLDPKWNOHYLDLYVG 63
Db 7 SRRIRVTIAADGLYKRDVFRFPDPFAVLTVDGE-QTHTTTAKKTLNPNYNNETFEVNT 65
QY 64 KTDSTISWNNHKKIHKQAGACFLGCVLLLSNAISRLLKDTGYQL--DLCKLNPSTDAV 121
Db 66 DNSTIAIQVFDOKKF-KKKGGQGLGVINLRVGVLDLALGDEMLTRDLKKN--ENTVV 122
QY 122 RGQIVVSLQ-----TRDRITGGSVVDCRGLLENGTVYED 157
Db 123 HGKIIINLSTTAQSTLQVPSAASGARTQRTSITNDPQSSQSSSVSRNPASSRAGSPTRD 182
QY 158 SGPG-----RPLSCFME-----PAPY-----TDSTG-----AAA 182
Db 183 NAPAASPASSPRFTSSFDQGLRPPGWERDNLGRYVYDHNTRSTTWIRNLSVA 242
QY 183 GGCNCRFVSPS-----QDRLQARLNPDPVRSGLQTPQNRPHGHQSPBLPGYEQ 234
Db 243 GAAAELHSSASSANVTGQVPSNAARTEASVLTSNATTAG-----SGELPPGWEQ 296
QY 235 RTTVQGVYVLTQTVSTVHDPRI PRDLNSV-----CDELGPLPGWEVRS 283
Db 297 RYTPGRIYFVDHNTRITTTVDPRQYIRSYGGPNNTIQQQVPSQGLPSPGWEMLT 356
QY 284 VSGRIYFVDHNTRITTTVDPRQYIRSYGGPNNTIQQQVPSQGLPSPGWEMLT 343
Db 357 NTARVYFVDHNTRITTTVDPRQYIRSYGGPNNTIQQQVPSQGLPSPGWEMLT 392
QY 344 LVOKLVLRHLSLQOQAGHCRIEVSREIFBESYRQIMKMPKDLKKRLMKVFRGEG 403
Db 393 FRKLIYFSLQAL-HPLPGQCHIKVRNHI FEDSYABIMRQSATDILKKRLMKVFRGEG 451
QY 404 LDYGGVAREWLYLICHEMLNPYYGLFOYSTDNIMQLQINPDSSINPDHLSYFHFVGRIMG 463
Db 452 LDYGLSREYFLLSHMFNPFYCLFEYSVDNYTLQINPHSGINPEHLNFKFGRVIG 511
QY 464 LAVFHGYINGGTVVPFKYKOLLGKPIQLSDLESVDPELHKSILWILENDITPVLDHTFCV 523
Db 512 LAIFHRFVDFAFFVYFVKMLQKVTLDQWESDAEYVESLWILNDITGVLDLTFV 571
QY 524 EHNAPGRILQHELKPNRNPVTENKEEYVRLYVNRFRMRGIEAOPFLAQKGNELIPQ 583
Db 572 EDNCFGEVITDLKPNRNIETVENKEEYVRLYVNRFRMRGIEAOPFLAQKGNELIPQ 630
QY 584 HLLKPPDQKELIIGGLDKIDLNDKSNTRKHKVADSNIVRWFQAVETFDERRARL 643
Db 631 ELINVFDERELELLIGGSEIDMEDWKKHDKYRSYENDQIKWFLMDWSENEKKSRL 690
QY 644 LQFTVGTSTVPLQGFALQGSTGAAGPRLFTIHLIDANTNLPAHTCFNRDIPPYESY 703
Db 691 LQFTGTSTVPLQGFALQGSTGAAGPRLFTIHLIDANTNLPAHTCFNRDIPPYESY 746
QY 704 EKLKLTAVETTCGFAVE 723
Db 747 KDLHKLTAIVETTCGFAVE 766
RESULT 3

QY	9	LTVLCAKNI	AKKDFRL	DDPFAKIVVDGSGQCHSTDTVKTALDPKMNCHYDLY	-----V	62
Db	9	LPVIAADG	LYKRDVFRFPDPFAVATINGE-OTKTTQVSKRTLPNPMWNEHDFWLIQRSKV	67		
QY	63	GKTDSTIT	VMNHHKIKKOGAGFLGCVRL--LSNAISRLKDTGVQRILDLCKLNPSTDAV	121		
Db	68	NEDSLAVQ	VFQKXP-KKQCGFLGVINVRIGVIELAPDAEDQMLRDLKKSTONLVV	126		
QY	122	RQOIVVS	QTR-----DRIGTGSYVDCRGLLENEGTVYED-----SGP-----	160		
Db	127	HGKLIINT	SLNTLATMSRLGPPSSRSRLTPQSSVISNDRANERPSSAMSGPENGANN	186		
QY	161	-----GRPL	SCFMEE-----PA-----PYTDTGT-----	179		
Db	187	MTLASRPAS	LAVSSSSTAPPTGTGTATPNSTVLVPAQARHSTLSPFEDSMGRLLPAGWE	246		
QY	180	-----	-----AAAGGNCRFVESPSQDRLQAQRLNPDPVRGSLQT	215		
Db	247	RREDHLGR	TYVDHNSRTTSMNRPTGTGAENRTAEANTQVERQBHQNRLLPDRTGANS	306		
QY	216	PQNRPH-	-----GHQSP---ELPEGYEQRTTVQGQVYFLHTQGTGVSTWHDPR	258		
Db	307	PLTQQQAAA	TANAAVMHGTGATPTGTGELPAGWEQRTPEGRPYFYVDHNTRTTTWDPR	366		
QY	259	I PRDL-	-----NSVN-----CDELGPLPGWEVRVTSGRIFYVDHNNRTTQFTDPL	305		
Db	367	RQOYRMYG	QNNNTGTTQQQFVSQLGPLSGEMWRLNTARIFYVDHNTKITTWDDPR-	425		
QY	306	HHIMHQOL	KEPSPOLPLPSEGSLEBELPAQYERDLVOKLURHELSLQOQOAGHC	365		
Db	426	-----	-----LPS--SL-DQNVDP--QYKEDFRKLIYFRSQPAM-RIMSGQC	461		
QY	366	RJESVREB	EIYESRQIMKMRPKOLKKRLMVKYPRGEGLDYGGVAREWLYLICHEMLNPY	425		
Db	462	HLKVRSHI	FEDSFABEISQATLKKRLMKFCEGDGLDYGUSREFFLLSHHEMFNF	521		
QY	426	YGLFOYST	NTIYMLQINPDSSINPDHLSYHFVGRIMGLAVFPHGYINGGTVFPFYKOLL	485		
Db	522	YCLFYFSA	HDNYTLQINPHSGINPEHLNYPKFIGRVVGLAIFHRRFLDAFFIGALYKMWL	581		
QY	486	GKPIGLDL	ESVDPELHKSLVMIENDITPVLQHTFCVEHNAEGRILQHELXPGENVVPY	545		
Db	582	GKAVSLAD	MEGVDAFDHRSLSQMLNDINDITDVLQATFTSEDEREGVITBEDLPINGENIAV	641		
QY	546	TEENKKEY	RVRLYVNNRFRMRGTEACFLAQKGFNELIPQHLKPFQDQKLELIIIGLDKID	605		

```
QY 183 -----GGGNCRFVE-----SPSQDRLQAQRLNPDVRSGLQTPQNR-----PHG 222
DB 234 GWEERTDNFGRTYVVDHNTTRTTWKPTLDQ--TEAERGQNLNANTELRQRGRGRTLPQG 292
QY 223 HQS-----PELPEGYEORTTVQGVYFLHT 247
DB 293 SSNSSVTVQGGSNIPVNGAAAAFAATGTTTSGLGELPSGWQRTTPGRAYFDH 352
QY 248 QTGVSTWHDPRIPRLDLSVN-----CDELGLPFGWEVRSTVSGRIYFVDHNNRT 297
DB 353 NRTTITWDRRQQYRTYGTPTNTTQQQPVSQLGLPSGWEMLNTARVYFVDHNTKT 412
QY 298 TQFTDPRLLHMHQCKLPSPQLPLPSGSELEDELPQRYERDLVQKLVRLHLSL 357
DB 413 TTWDDPR-----LPS--SL-DQNV--QYKDRFRKVIYFRSQPAL 448
QY 358 QPQAGHCRTEVSGRETEFESSYQIMQRPKDLKRLMKVFRGEEGLDYGVAREWLYLL 417
DB 449 -RLIPGOCHIKVRKNTIFDAYQEIWRQTPEDLKRLMIKFDGEEGLDYGVSREFFELL 507
QY 418 CHEMLNYYGLFOYSTDNIIYMLQINPDSSINPDHLSYFFHVGIMGLAVPHGYINGGT 477
DB 508 SHEMFNPFYCLFEYSAYDNTYTIQINPSGINPBLAYFKFGRVWGLGVFHRFLDAFV 567
QY 478 VPFYKQLGKPIQLSLESVDPELHKLSVLWLENDITPVLDTHTFCVEHNAFGRILQHELK 537
DB 568 GALYKMLRKKVVLQDMEGVDAEYVNSLWKLWLENSIDGVLDTLFSADDERFGEVTVDLK 627
QY 538 PNGRNVPVTEENKKEYVRLVYNNRFRMGIEBAQFLALQKGFNELIPQHLKPPQKLELI 597
DB 628 PDGNEIVTDCNKKKEYVELYVQWRIVDVQVQKAFMDGFENLIPEDLVTVFDERELELL 687
QY 598 IGGDLKIDLNDKSNTRLKHCVDADSNIVRMFWQAVETFDERRARLLQFVTGSTRVPLQG 657
DB 688 IGGFAEIDIEDWKHTDYRGQESDEVIOFWKVCSEWNEQEARLLQFTTGTSTRVPLNG 747
QY 658 FKALQGSTGAAGPLFTIHLIDANTDNLPAKHTCFNRIIDIPPYESYKLYEKLTAVEET 717
DB 748 FKDLQGSQD---GPRRFTIEKA-GEVQQLPKSHCTCFNVDLPQYVDYDSMKQKLTAVEET 803
QY 718 CGPAVE 723
DB 804 IGFQGE 809

RESULT 5
T39585
ubiquitin protein ligase - fission yeast (Schizosaccharomyces pombe)
C:Species: Schizosaccharomyces pombe
C>Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 18-Aug-2000
C:Accession: T39585
R:Volckaert, G.; Wood, V.; Rajandream, M.A.; Barrell, B.G.
submitted to the EMBL Data Library, August 1997
A:Reference number: Z21865
A:Accession: T39585
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-786 <VOL>
A:Cross-references: EMBL:Z99759; PIDN:CAB16903.1; GSPDB:GNO00067; SPDB:SPBC16E9.11c
A:Experimental source: Strain 972h; cosmid c16E9
C:Genetics:
A:Gene: SPDB:SPBC16E9.11c
A:Map position: 2
A:Introns: 60/2
C:Superfamily: Yeast ubiquitin-protein ligase; WW repeat homology
F:236-273/Domain: WW repeat homology <WW1>
F:306-343/Domain: WW repeat homology <WW2>
F:364-401/Domain: WW repeat homology <WW3>

Query Match 34.5%; Score 1341; DB 2; Length 786;
Best Local Similarity 36.6%; Pred. No. 2.1e-89;
Matches 304; Conservative 128; Mismatches 236; Indels 162; Gaps 21;
```

```
QY 6 KIRLTVLCAKNLAKOFFRLPDPAKIVVDGSGQCHSTDTVKNTLDPKMKHQHYDLYVGKT 65
DB 7 RVRVYVAAAGLSKRDLFRQDPDFALLTVDGE-QTHYTKVKKSVAPYWNESGEFVTVKPS 65
QY 66 DSITISVMNHKKHKHKGAGAGLCVRLLSNAISRLKDTGYQR-----LDLCK-LNPSDTD- 119
DB 66 SVISIRLFDQKF-KKXQOGLGLVSP-----RMREVGSFRSNREVSLTEP-LKKSSTTN 118
QY 120 -AVGQIVW-----SL 129
DB 119 LSVLGNLVLKVAPSKIRAPACNHSSTTANRTTPTTTTARTTTPRPTATNTNQST 178
QY 130 QTRDRIGTGSVVDRCGLLENEGTVYEDSGPGRPL-----SCFMEEPA-----PYTDS 177
DB 179 SNSITENGTSAASTNGTGTGAGTGASHSS-----PVTNRQTNNTSALSNNAHIMSSPEDQ 234
QY 178 TGAAGG-----GNCRFVE-----SPSQDRLQAQRLNPDV 209
DB 235 YGRLLPPGWERRADSLGRYYVDHNTTRTTWTRPASSNTPVHNTSSDSQRLNHQNRHLPD- 293
QY 210 RGSLLQTPQNRPHGHQSP---ELPEGYEORTTVQGVYFLHTQTGVSTWHDPRIP- 260
DB 294 -----DSNPSLMQSDSGNDLPGWEVRYTDTGCPYFVDHNTTITWVDPNPLVRPNG 346
QY 261 -----RDLNSVNCDELGLPPLPGWEVRSTVSGRIYFVDHNNRTTQFTDPRLLHMHQCCQ 314
DB 347 GSSTVGSMLQPQSLSHLGLPLSGWEMBLTNSARYFVDHNTKTTTWDPR- 397
QY 315 LKESQPLPLSEGLSEDELPQRYERDLVQKLVRLHLSLQPOAGHCRIEVSRREI 374
DB 398 -----PSALDQDVP--QYKCDFRKLIYFRSQPM-RPLFGQCNVVRDHI 441
QY 375 FEESYRQIMKMPKDLKRLMKVFRGEEGLDYGVAREWLYLLCHEMLNPPYGFQISTD 434
DB 442 FEDSYAEIMRYSAHDLKRLMIRFDGEDGLDYGGLSREFFLLSHKMFDPYICLFYSAV 501
QY 435 NIYMLQINPDSSINPDHLSYFFHVGIMGLAVPHGYINGGTVPYFKQLGKPIQLSDL 494
DB 502 DNYTQINPHSINPEHLNRYFRIGRVIGLAIHRRFLDAFFVSLYKLLRKVKVSLADM 561
QY 495 ESDVDPELHKS-VWILENDITPVLDTHTFCVEHNAFGRILQHELKPNRNVPTTEENKKEYV 554
DB 562 ESIDAEFYRSLKWLLENDITGLDITFSVEEDHGEVTRVELITNGENIEVTEENKKEYV 621
QY 555 RLYVNRVWRMGIEAQFLALQKGFNELIPQHLKPDQKELELIGGLDKIDLNDKSNTR 614
DB 622 DLVTEWRYSKRVEQQFNAFYSGFVELVSPDLVNVDFERELELLIGGISDVEDVWKSHTE 681
QY 615 LKHCVDADSNIVRMFWQAVETFDERRARLLQFVTGSTRVPLQGFKALQGSTGAAGPLFT 674
DB 682 YRTYIATDPVIMKFWELIAGWKQEDRSKLLQFAIGTSIIPVNGFEDLQGSQD---GPRKFT 738
QY 675 IHLIDANT-DNLPAKHTCFNRIIDIPPYESYKLYEKLTAVEETCGPAVE 723
DB 739 IE-XAGTPDQLPVAHTCFNRLDLPDYPKDTLHEKLSLAVENTVGFQNE 786

RESULT 6
T46412
ubiquitin-protein ligase (EC 6.3.2.19) NEED4 - human (fragment)
N:Alternate names: hypothetical protein DKFZp434P2422.1
C:Species: Homo sapiens (man)
C>Date: 04-Feb-2000 #sequence_revision 04-Feb-2000 #text_change 03-Jun-2002
C:Accession: T46412
R:Blum, H.; Bauseachs, S.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.
submitted to the Protein Sequence Database, January 2000
A:Reference number: Z23034
A:Accession: T46412
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-820 <AAA>
A:Cross-references: EMBL:AL1137469
```

A:Experimental source: adult testis; clone DKFZp434P2422

C:Genetics:

A:Gene: GDB:NEDD4

A:Note: DKFZp434P2422.1

C:Superfamily: rat ubiquitin-protein ligase; protein kinase C C2 region homology; ubiquitin

C:Keywords: ligase

F:142-179/Domain: WW repeat homology <WWR1>

F:342-379/Domain: WW repeat homology <WWR2>

F:393-430/Domain: WW repeat homology <WWR3>

F:489-814/Domain: ubiquitin-protein ligase homology <UBI>

Query Match 32.0%; Score 1241.5; DB 2; Length 820;

Best Local Similarity 34.9%; Pred. No. 3.9e-82;

Matches 294; Conservative 116; Mismatches 230; Indels 203; Gaps 21;

QY 43 TDTVKTLDPKNQHVDLVVGTDT-SITISVNNHKKHKKQAGAGLGCVRLLSNAL----- 97

DB 12 TTKIITKLTPKNNEEFYFVNFSNRLLEVEVDENLRITDD---FLGQVDVPLSLHPTED 68

QY 98 -----SLKDTGYQRLDLCKLNFSDTDVAVRGQIVSVLSQTRDRIQT 137

DB 69 PTWERPYTFKDFLLRPSRSHKSVK--GFLRLKWAYMPKNG-----GQDEENSQDRDMEH 121

QY 138 GGSVVDRCG-----LLENEGTVY----- 155

DB 122 GMEVVDNSDSASQHOELPPPLPGWEEKVDNLGRTYYVNNHNRRTTQWHRPRLMDVSS 181

QY 156 -----EDSGP-----GRPLSCFMEEPAPYTDSTGAA-- 181

DB 182 SDNNIRQINQEAARRFRSRHISEDLPEPSEGQDVPEPWETISEEVNIAGDSLGLALP 241

QY 182 ---AGGNCRFVESPSQD-----ORLQAQRLN----- 206

DB 242 PPPASFGSTSPQELSELSRLQITPDSNGQFSLSIQREPSSRLSCVSTDAVAGQGH 301

QY 207 -----PDVRGSLQTPNRRPHGHOSPE-----LPGEYEQRTTVQGVFL 245

DB 302 LPFGAKDSPRAVADTSLNPSQSPSPYNSPKQHKVQSLQPLPGCWENRIAPNGRPFI 361

QY 246 HTQTGVSTWHDRIIP-----RDLSVNCDELGLPPGWEVSTVSGRIYFVDHNNRITQF 300

DB 362 DHTKTTTWTWEDRLKFPVPMHRSKTSLNPNLGLPPGWEERIHLDRGTFYIDHNSKIQT 421

QY 301 TDPRLHHIMHOCQLKEPSQPLPSEGSLEDEELPAQRYERDLVOKLVLRHLSLQCP 360

DB 422 EDPRQN-----PAITG-----PAVPYGRFKQYDYFRKKLKKPAD 458

QY 361 QAGHCRIEVSREIEEESYRQIMKM-RPKDKLKKLWVFRGEGLDYGGVAREWLYLLCH 419

DB 459 IPNRPEMLKLNHNIIEESYRIRMSVKRPDVLKARLWIEFESEKGLDYGVAREWFLLSK 518

QY 420 EMLNPPYGLFQYSTNIMYLNQINPDSSI-NPDHLSYFHFVGRIMGLAVFHGHYINGGTV 478

DB 519 EMFNPPYGLFEYSATDNTYLTQINPNSGLCNEHDHLSYFTFICKVAGLAVFHGKLDDGFFIR 578

QY 479 PFYKOLLGKPIQLSDLESVDPLHLSLVILENDITPVLDITFCVEHNAFGRILQHELKP 538

DB 579 PFYKMLGKQITLNDMESVDSEYNSLKWILENDPTE-LDLMFCIDENFQGTQVQDLKP 637

QY 539 NGRNVPTVEENKKEYVRLVNNRFRGIEAQFALQKGFNELIPOHLLKPFDOKELELI 598

DB 638 NGEIMVTNENKREVIDLVIQWRVNVKQMNAPLEGFTTELLPIDLIKIFDENELELLM 697

QY 599 GGLDKIDLNDKSNRLKXHCVADSN-IVRWFWQAVETEDERRARLLQFVTGSTRVPLOG 657

DB 698 CGLGDVDNDWQHSHYKNGYCPNRPVQWFWKAVLLMDASKIRLLQFVTGSRVPMNG 757

QY 658 FXALOGSTGAAGPRLFTIHLIDANTDNLKFAHTCFNRIDIPPESEYKLYEKLTAABET 717

DB 758 FAELYGSN--GPQLFTIEQW-GSEPKLPRAHTCFNRDLDPPEYTFEDLREKLLMAVENA 813

QY 718 CGF 720

DB 814 QGF 816

RESULT 7

S70642

ubiquitin ligase Nedd4 - rat (fragment)

C:Species: Rattus norvegicus (Norway rat)

C:Date: 14-Feb-1997 #sequence_revision 13-Mar-1997 #text_change 18-Aug-2000

C:Accession: S70642

R:Staub, O.; Dho, S.; Henry, P.C.; Correa, J.; Ishikawa, T.; McGlade, J.; Rotin, D.

EMBO J. 15, 2371-2380, 1996

A:Title: WW domains of Nedd4 bind to the proline-rich PY motifs in the epithelial Na(+)

A:Reference number: S70642; MUID:96221297; PMID:8665844

A:Accession: S70642

A:Status: preliminary; nucleic acid sequence not shown

A:Molecule type: mRNA

A:Residues: 1-887 <STA>

A:Cross-references: EMBL:U50842; NID:G1293646; PIDN:AAB48949.1; PID:G1293647

C:Genetics:

A:Gene: Nedd4

C:Superfamily: rat ubiquitin-protein ligase; protein kinase C C2 region homology; ubiquitin

F:54-167/Domain: proline kinase C C2 region homology <KC2>

F:246-283/Domain: WW repeat homology <WWR1>

F:402-439/Domain: WW repeat homology <WWR2>

F:459-496/Domain: WW repeat homology <WWR3>

F:555-881/Domain: ubiquitin-protein ligase homology <UBI>

Query Match 31.6%; Score 1227.5; DB 2; Length 887;

Best Local Similarity 33.9%; Pred. No. 4.6e-81;

Matches 296; Conservative 127; Mismatches 224; Indels 227; Gaps 28;

QY 7 IRUTVLCANLAKKOFFRLEPDPFAKIV---VDGSGQCHSTDTVNTLDPKN----- 55

DB 77 VRVKVIAGIGLAKKILGASDPYRVTVLYDPMGVLTSVQTKIKSLNPKNWEILFRV 136

QY 56 ---QH-----YDLYVQKTD--SITISVNN-----HKIKHKQG 83

DB 137 LPQHRILFEVDENRUTRDDFLGQVDVPLYPTENPRMERPYTFKDFVLHFRSHKSRV 196

QY 84 AGELGCVRLLSNAISRLKDT-----RLKMTYLPKNGSDDENADCAEELEPGWVVLQDPDAATHLQHPPEP 245

DB 197 KGYL-----KX-QLDLG-----KLNPSD--TDVVGQIVV---SLQTRDRI 135

QY 246 SPLPPGWEERQDVLGRYYVNHESRTTQMKRPSDEDLTDDENGDIQLQAHGAFTTRQI 305

QY 136 GTGSGVVDRCG-----LLENEGTVYEDSGFRPLS-----C 166

DB 306 SE-----DVGDPDNHSPENWEIVREDNTIYSGAVQSPSPGHDPVQVRLABELDTRLT 360

QY 167 FMEIPA---PYTDSGTAAAGGN--CRFVESPSQDQRLQARLNPDVGRSLQTPQNRPH 221

DB 361 MYGNPATSQPVTSNNHSSRGSGTCTFEEQPTLPVLLPT----- 400

QY 222 GHOSPELPEGYEORTTVQGVYFLHTQTGVSTWHDH-----RIPRDLNS---VNCDE 270

DB 401 ---SSGLPPGWEERQDGRSYIVDHNSTKTTTWSKPTMQDDPRSKIPAHLRGKTPTVDSND 457

QY 271 LGPLPPGWEVRSVSGRIYFVDHNNRRTTQFTDPRLLHHIMHOCQLKEPSQPLPSEGSL 330

DB 458 LGPLPPGWEERTHDCGRVFINNIKTQWEDPRMQNV-----AITGPAEP----- 503

QY 331 EDEELPAQRYERDLVOKLVLRHLSLQPOQACHCIEVSREIEFESYRQIMKRPKD- 389

DB 504 -----YSRDYKRYEFREKLLKQKQTDIPNKFEMKLRANILEDSYRRIMGVRRADF 554

QY 390 LKRLMKVFRGEGLDYGGVAREWLYLLCHEMLNPPYGLFQYS-TDNIMYLNQINPDSSI- 447

DB 555 LKARLWIEFGEKGLDYGGVAREWFLISKEMENPYGLFEYSATEDNTYLTQINPNSGLC 614

QY 448 NPDHLSYFHFVGRIMGLAVFHGHYINGGTVVPYKOLLGKPIQLSDLESVDPLHLSLVW 507

```
Db 615 NEDHLSYFKPIGRVAGMAVYHGLDGLDFFIRPFYKMKLQKLIHLTHDMESVDSEYSSSLAW 674
QY 508 ILNDITPVLDTFCVHNHAFGRILQHELKFNCRNVPVTEENKEVYVLYVNRFRMRGIE 567
Db 675 ILNDPTE-LDLRPIDEELFGTHQHELKGTGSEVVVTKNKKKEIYLVIOQRVNRIO 733
QY 568 AQFLALQKGFNELIPQHLKPPQKLELIGGLDKIDLNDKSNRRLKHCHA-DSNIVR 626
Db 734 KQMAAFKRGFFELIPQDLIKIFDENELLMGLGDVDVNDWREHTKYNGYSLNHQVTH 793
QY 627 WFWQAVETFEERRARILQVTCSTVPLOGFKALOGSICGAGPRLFTIHLIDANTDNL 686
Db 794 WFWKAVLMDSEKRIQLQVTCSTVPVMMGFALYGSN---GPGSFTVEQW-GTPDKLP 849
QY 687 KAHTCFNRIDIPPVSEYKELYKLLTAVERTCGF 720
Db 850 RAHTCFNELDLPVYESFDELWDLQMAIENTQGF 883

RESULT 8
NEDD-4 ORF - mouse (fragment)
C:Species: Mus musculus (house mouse)
C:Date: 02-Aug-1996 #sequence_revision 02-Aug-1996 #text_change 18-Aug-2000
C:Accession: 183196
R:Kumar, S.; Tomooka, Y.; Noda, M.
Biochem. Biophys. Res. Commun. 185, 1155-1161, 1992
A:Title: Identification of a set of genes with developmentally down-regulated expression
A:Reference number: I60167; MUID:92328780; PMID:1378265
A:Accession: 183196
A:Status: Preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-708 <RES>
A:Cross-references: GB:D10714; NID:g220508; PID:g220509
C:Genetics:
A:Gene: NEDD-4
C:Superfamily: rat ubiquitin-protein ligase; protein kinase C C2 region homology; ubiquitin
F:40-77/Domain: WW repeat homology <WW1>
F:196-233/Domain: WW repeat homology <WW2>
F:251-288/Domain: WW repeat homology <WW3>
F:347-682/Domain: ubiquitin-protein ligase homology <UBI>

Query Match 27.1%; Score 1051.5; DB 2; Length 708;
Best Local Similarity 36.1%; Pred. No. 2.2e-68;
Matches 244; Conservative 102; Mismatches 194; Indels 135; Gaps 19;

QY 50 LDPKWNQHYDYVGKTSITISVWNHKKHKKQAGLGCVRLLSNRAISLKDGTGYRLD 109
Db 42 LPPGWEERQDV-LGRITYV---NHES-----RRTQWKR-- 70

QY 110 LCKLNP-----DIDAVRGQIVSLQTRDRIGTGGSVVDCRGLLEN-----EGTV 154
Db 71 -----PSFDDDLTDNDMDQLQRAFTTRQISEVDGPDNRSENWNIIVREDENTE 125

QY 155 YEDSGPGRPLSCFME-----EPA---PYTDTGAAAGG--NCRPVES 192
Db 126 YSGQAVQSPSGHDIDVQTHLAEEFNTRLAVGCNPNATSPQVTSNHSRSGSLQTCIFEEQ 185

QY 193 PSQORLQAOQLRNPVDRGSLQTPQRPHGHSQBELPEGVEORTTVQGVYVLTHTGTGS 252
Db 186 PTLPVLLPT-----SSGLPPGWEKQDDRGRIYVDHNSKTI 222

QY 253 TWHDP-----RIPRDL-NSVNCDELGPLPGWEVRSTVSGRIYFVDHNNRTQTDP 303
Db 223 TWSKFTMQDDPRSKIPAHRLRGKTDNDLGPLPGWEERTHTDGRVFTINHNKKTOWEDP 282

QY 304 RLHIMNHQCLKEPSQPLPLPSGSLDEBELPAQRVERDLVQKLKVLRLHLSLQOQPAQ 363
Db 283 RLQNY-----AITGPAV-----YSRDYKRYEFRRKLKKQTQDIPN 319

QY 364 HCRTEVSREEIFESYROIKMKRPD-LKKELVMKFRGEGLDYGGVAREWLKLLCHEML 422
Db 320 KFEKLRRLANILEDYRIIMGVKGADLLKALWIEFDGKGLDYGGVAREWFFLISKEMF 379
```

```
QY 423 NPYVGLFQYSTDNTYMLQINPDSI-NPDHLSYFHFVGRIMGLAVFHGHYINGGFTVPFY 481
Db 380 NPYVGLFQYSDNTYMLQINPDSI-NPDHLSYFHFVGRIMGLAVFHGHYINGGFTVPFY 439
QY 482 KQLGKPTQLSDLESVDPELHKSILVLENDITPVLQHTFCVEHNAFGRILQHELKPNGR 541
Db 440 KQLGKPTQLSDLESVDPELHKSILVLENDITPVLQHTFCVEHNAFGRILQHELKPNGR 498
QY 542 NVPVTEENKKEVYVLYVNRFRMRGIEAQFLALQKGFNELIPQHLKPPQKLELIGGL 601
Db 499 EIVVTNKNKKKEIYLVIOQRVNRIOKQMAAFKRGFFELIPQDLIKIFDENELLMGL 558
QY 602 DKIDLNDKSNRRLKHCHA-DSNIVRWFQAVETFEERRARILQVTCSTVPLOGFKA 660
Db 559 GDVDVNDWREHTKYNGYSNMHQVHFWKAVMMDSKRIQLQVTCSTVPVMMGFAL 618
QY 661 LOGSTGAAGPRLFTI 675
Db 619 LYGSN---GPGSFTV 630

RESULT 9
T37900
probable ubiquitin-protein ligase - fission yeast (Schizosaccharomyces pombe)
C:Species: Schizosaccharomyces pombe
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 18-Aug-2000
C:Accession: T37900
R:Rieger, M.; McDougall, R.C.; Rajadream, M.A.; Barrell, B.G.
submitted to the EMBL Data Library, September 1999
A:Reference number: Z21752
A:Accession: T37900
A:Status: Preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-671 <R1E>
A:Cross-references: EMBL:AL117390; PIN:CA55856.1; GSPDB:GN00066; SPDB:SPAC1805.15C
A:Experimental source: strain 972h-; cosmid c1805
C:Genetics:
A:Gene: SPDB:SPAC1805.15C
A:Map position: 1
A:Introns: 80/2; 105/1; 639/2
C:Superfamily: WW repeat homology
F:242-279/Domain: WW repeat homology <WWR>

Query Match 26.4%; Score 1026.5; DB 2; Length 671;
Best Local Similarity 32.9%; Pred. No. 1.3e-66;
Matches 248; Conservative 128; Mismatches 253; Indels 125; Gaps 18;

QY 6 KIRLTVLCAKMLAKDFFRLDPPEAKIVVDGSGGCHSDTIVKNTLDPKWCHYDLYVGKT 65
Db 7 EQLVTILHVEGLWKNLRLSLKPYLLISVD-DDOFIKTNVASGTLRLSWGFTQKLTVPQ 65

QY 66 DSITTSVWNHKKHKKQAGLGCVRLLSNA---ISRLKDTGYQRLDICKLNPSTDPAVR 122
Db 66 STILLQLFDEKQ-KNETSDGFGVLGAADVNSFLPNPFKDDYKTRITL---RSPSGSYR 120

QY 123 GQIVVSLQTRDRIGTGGSVVDCRGLLENEGTVEDSGPRLSCFMEBPAP-----YTD 176
Db 121 GSVWCLFKR-----SKFLPEELPADKSIQCTD 147

QY 177 STGAAG-----GNCRFVESP---SQDORLQAOQLRNPVDRGSLQTPQRPHGHQS 225
Db 148 IDDSAGCAWETRIDEFQHYVLYKSPQLSVISASISHELEN-----LTFK----- 192

QY 226 PELPEGYQRTTVQGVYVLTHTGTGVTW-----HDPRIPLDNLNSVNCDEL 271
Db 193 -QLKEVFSQ-----FLFNQSKSLKINLEYKVIKHLLEHYF-LALSVRQOQVAVEK 241

QY 272 GPLPQGWVRSVTSGRIFYVDHNNRTQTTPRLHLMNHQCLKEPSQPLPLPSGSL 331
Db 242 GPLPQGWVRSVTSGRIFYVDHNNRTQTTPRLHLMNHQCLKEPSQPLPLPSGSL 284

QY 332 DBELPAQRVERDLVQKLKVL--RHLSLQOQPAQHCRIEVSREEIFESYROIKMKRPD 389
```



```
Db 295 ENTDSIQINDEYQRIAMVDRPEMAVNDQA---LQIKVRATTFEDAYDIISKLSVSD 341
Qy 390 LKRLMWFRGEGDYGVAEWYLLCHEMLNPPYGLFOYSTDNIMQLQINPDSSINP 449
Db 342 MKKLLIRPNEDGLDYGVSSEFFYLISHAIFNPGYSIFVATDNDVGLQISPLSSVNP 401
Qy 450 DHLSTFHFVGRIMGLAVFHGHVINGGFTVPFYKQLLGRPIQLSDLESVDPELHKSILWIL 509
Db 402 DFRSYFRFVGRVWGLAIYHRRYLDVQVLPFYKRIQLKPLCLDVKVDVEYVESLKWIK 461
Qy 510 ENDITPVLDTFCVCHNAFGRILQHELKPNGRNVPVTENKEKYEYVRLVYVNRFRMGIEAQ 569
Db 462 NNDVDESCLNFSVENRFGESVTDLPNGRNIAVNNQKNYLLKALTEHKLVTSTBEQ 521
Qy 570 FLALQGFNELIPQHLKPFDDQKLELIGGLDKIDLNDWKSNTLKHCVADSNIVRWFW 629
Db 522 FNALAGGLNELIPDSVLQIFENELDTLNGKRDIDVDQDKRFTDYRSYTTDDIVIFW 581
Qy 630 QAVENFDEERRARLLQFVTGSTRVPLQGFKALQSGTGAAGPRLFTIHLIDANTDNLPAKH 689
Db 582 EULSWSPKKAKLLQFATGTRSLPSLGFQVHGSD---GPKRFTIEKV-GHISQLPAKH 637
Qy 690 TCFNRIDIPPYSEYKLYEKLKTAVEETCGFAVE 723
Db 638 TCFNRIDIPPYNSKEEQLKLTIAIQTAGFGE 671

RESULT 10
T49799
N:Alternate names: protein [imported] - Neurospora crassa
C:Species: Neurospora crassa
C:Date: 02-Jun-2000 #sequence_revision 02-Jun-2000 #text_change 02-Jun-2000
C:Accession: T49799
R:Schulte, U.; Aign, V.; Hohelsel, J.; Brandt, P.; Fartmann, B.; Holland, R.; Nyakatura,
submitted to the Protein Sequence Database, May 2000
A:Reference number: Z25022
A:Accession: T49799
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-3839 <SCH>
A:Cross-references: EMBL:AL356834; GSPDB:GN00116; NCSP:B11B22.10
A:Experimental source: BAC clone B11B22; strain OR74A
C:Genetics:
A:Gene: NCSP:B11B22.10
A:Map position: 6
A:Intons: 16/3; 2607/1; 2623/1; 2658/1; 2987/2; 3204/3; 3694/1; 3809/1

Query Match 22.0%; Score 854.5; DB 2; Length 3839;
Best Local Similarity 39.4%; Pred. No. 6.6e-53;
Matches 178; Conservative 92; Mismatches 155; Indels 27; Gaps 8;

Qy 275 PRCWEVRSTVSGRIY-FVDHNRRTTQTPRLHHIMHQCQLKEPSQPLPLPSEGSLEDE 333
Db 3410 PP-----EDRIAGLFFTFTEHRR-----LNEVLRRHNPKLMSGTFSLVKNPKVLEFD 3458
Qy 334 ELPAQYERDLVOKLKVLRHLSLQOQAGHCRIEVSREEIFEESYRQIMKMRPKDLK-K 392
Db 3459 N-KRNVFNSVHSKYQOTRHSPPQLQ-----LQVRREHVFHDSFRSLYKAKDELKFG 3510
Qy 393 RLMVKFRGEGDYGVAEWYLLCHEMLNPPYGLFOYSTDNIMQLQINPDSSINPDL 452
Db 3511 KLNIRFCGEGVDAGGVTRFEWFOVLRQMPDNVYLVFVPSVSDRTTFHKNKLSPINDEHL 3570
Qy 453 SYFHFVGRIMGLAVFHGHVINGGFTVPFYKQLLGRPIQLSDLESVDPELHKSILWILEND 512
Db 3571 PPFKPTGRIIGKALYGRLLCEYFSRAVYKRIIGKPVSKDMESFPDYYKSLVWLENLND 3630
Qy 513 ITPVLDTFCVCHNAFGRILQHELKPNGRNVPVTENKEKYEYVRLVYVNRFRMGIEAQFLA 572
Db 3631 ITDITETESVEDVFGVKVVDLIENGRNIPVTENKEHYVRLIVEHKLITTSVKDQMKKA 3690
```

```
Qy 573 LQKGFNELIPQHLKPFDDQKLELIGGLDKIDLNDWKSNTLKHCVADSNIVRWFWQAV 632
Db 3691 FLTGFEHIIPEELIAIFNEQLELLISGUPDIDIDWKANTYHNSAGAPQIQWFRAV 3750
Qy 633 ETDEERRARLLQFVTGSTRVPLQGFKALQSGTGAAGPRLFTIHLIDANTDNLPAKHTCF 692
Db 3751 RSDKEELAKLQFVTGSTRVPLNGFKLEGMGVS---RFNIHRDYGSKORLPSSHTCF 3807
Qy 693 NRIDIPPYSEYKLYEKLKTAVEETCGFA 721
Db 3808 NOLDPEYENYETLRSQLLKAITAGSDYFGFA 3839

RESULT 11
T37964
probable ubiquitin ligase - fission yeast (Schizosaccharomyces pombe)
C:Species: Schizosaccharomyces pombe
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 03-Dec-1999
C:Accession: T37964
R:Devlin, K.; Churcher, C.M.; Wood, V.; Bartelli, B.G.; Rajandream, M.A.
submitted to the EMBL Data Library, September 1997
A:Reference number: Z21757
A:Accession: T37964
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-3227 <DEV>
A:Cross-references: EMBL:Z99531; PIDN:CAB16714.1; GSPDB:GN00066; SPDB:SPAC19D5.04
A:Experimental source: strain 972h-; cosmid C19D5
C:Genetics:
A:Gene: SPDB:SPAC19D5.04
A:Map position: 1

Query Match 21.8%; Score 848; DB 2; Length 3227;
Best Local Similarity 43.1%; Pred. No. 1.5e-52;
Matches 170; Conservative 79; Mismatches 125; Indels 20; Gaps 5;

Qy 344 LVQKLKVLK-----HLSLQOQAGHCRIEVSREEIFEESYRQIMKMRPKDLK 391
Db 2838 LVKNPKVLEPENKENYFNRLHEEAAKE-QYPLNITVRDHFVLDYSRALHFKDADEVK 2896
Qy 392 -KRLMWFRGEGDYGVAEWYLLCHEMLNPPYGLFOYSTDNIMQLQINPDSSINPD 450
Db 2897 FSKLNIHFRDEEGVDAGVTRWLOVLARQVFPDYALFLPVTGDATTFHPNRDSSVNP 2956
Qy 451 HLSYFHFVGRIMGLAVFHGHVINGGFTVPFYKQLLGRPIQLSDLESVDPELHKSILWILE 510
Db 2957 HLSFFKTRIIGKALYGRLLDCHFRSAVYKMLHRSVSVKDIESLDPPYKSLVWMLN 3016
Qy 511 NDITPVLDTFCVCHNAFGRILQHELKPNGRNVPVTENKEKYEYVRLVYVNRFRMGIEAQF 570
Db 3017 NDITDITTEFAFAKDVFGKTVVDLLIPNGRNIPVTENKQNVYVNRWMDIKLRESVKDQL 3076
Qy 571 LALQKGFNELIPQHLKPFDDQKLELIGGLDKIDLNDWKSNTLKHCVADSNIVRWFWQ 630
Db 3077 KSLLDGFSDIIPSHLIQIFNEQELELLISGLPEIDIDDWKNNTYHGYNVSSPQVQFWFR 3136
Qy 631 AVETFDERRARLLQFVTGSTRVPLQGFKALQSGTGAAGPRLFTIHLIDANTDNLPAKHT 690
Db 3137 AVRSFDEERAKLQFATGTSKVLNFKLEGMGSGF---QRFNIHKSYSGLNRLPOSHT 3193
Qy 691 CFNRIDIPPYSEYKLYEKLKTAVEETC---GFA 721
Db 3194 CFNQLDPEYDVTYQLRSMLLTAINESGEGFGFA 3227

RESULT 12
H96599
protein F14J16.10 [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
C:Accession: H96599
R:Theologis, A.; Becker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;
```

ansen, N.F.; Hughes, B.; Huizar, L.
 Nature 408, 816-820, 2000
 A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
 C:Authors: Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziani,
 Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
 A:Authors: Salzbeg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
 Ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
 A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
 A:Reference number: AB6141; MUID:21016719; PMID:11130712
 A:Accession: H96599
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-4056 <STO>
 A:Cross-references: GB:AE005173; NID:98778329; PIDN:AAF79338.1; GSPDB:GN00141
 C:Genetics:
 A:Gene: F14U16.10
 A:Map position: 1

Query Match 20.2%; Score 784; DB 2; Length 4056;
 Best Local Similarity 39.4%; Pred. No. 9.8e-48;
 Matches 172; Conservative 65; Mismatches 138; Indels 62; Gaps 7;
 A:Map position: 1

QY 343 DLVQKLVLRHLSLQPPQ--AGHCRIEVSREEIFEESYQIMKMRPKDLKKRLMVKFRG 400
 DB 3624 DFDNKKAYFRSRIRHQDHISGRLISVRRAYVLEDSYNQLRMRSPODLKGRNLNVQFQG 3683
 QY 401 BEGLDYGGVAREWMLYLICHEMLNPPYGLFOYSTDNIMYLQINPDSSINPDHLSYFHFVGR 450
 DB 3684 BEGDAGLTREWYQLLSRVIFDKGALLFT-TVGNDAFQFNPNNSVYQTEHLSYFKFVGR 3742
 QY 461 IMGLAVFHGHYINGGFTVPFYKQLLKPQLSDLESVDPELHKLWILENDITPVLDT 520
 DB 3743 MVAKALPDQLLDVYFTRSPYKHLGVKTVHDIEAVDPDYKXKLWLENDVSDILDT 3802
 QY 521 F-----CQE-----HNAFORILQHELKPNGRN 542
 DB 3803 FMSDADEKHILYKTEVRLMFCFPLFWCFIPKCHCIEIILSLMKKVTYDELKPGGRN 3862
 QY 543 VPTVEENKKEYVLYVNWFRMRGTEAQLALOKGFNELIPOHLKPPDKOKELELIIGGLD 602
 DB 3863 IRVTEETKHEYVDLVAGHILTNAEPQINAFLEGFNELIPRELVSINDKELELLISGLP 3922
 QY 603 KID-----LNDWKNTRLKHCVADSNIVRWFQAVETFDERRARLLQFV 647
 DB 3923 EIDCKLSDIQYLCAYPDLIDLKANTETSYTAGSPVIRWFWEVVKAFSKEDMARFLQFV 3982
 QY 648 TGSTRVPLQGFALQGGTGAAGRLFTHLIDANTDLNPKAHTCFNIDIPPYESYEKLY 707
 DB 3993 TGTSEKVLPEGFALQ---GISGPORLQHKAYGAPERLPSAHTCFNQLDLPFYQSKEQLQ 4039
 QY 708 EKLTAVEETC---GFA 721
 DB 4040 ERLLLATHEASEGGFGFA 4056

RESULT 13
 ubiquitin-protein ligase homolog F1707.15 - Arabidopsis thaliana
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C:Date: 12-Feb-1999 #sequence_revision 12-Feb-1999 #text_change 04-Mar-2000
 A:Accession: T01491
 R:Vysotskaya, V.S.; Schwartz, J.R.; Toriumi, M.; Yu, G.; Kwan, A.; Oji, O.; Liu, S.; Li,
 xtz, D.; Li, Y.; Palm, C.J.; Shinn, P.; Sun, H.; Davis, R.W.; Ecker, J.R.; Federspiel, N.
 submitted to the EMBL Data Library, June 1998
 A:Description: Arabidopsis thaliana chromosome 1 BAC F1707 sequence.
 A:Reference number: Z14334
 A:Accession: T01491
 A:Status: translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-1126 <YS>
 A:Cross-references: EMBL:AC003671; NID:92833627; PID:G3176690; GSPDB:GN00059; ATSP:F1707
 A:Experimental source: cultivar Columbia
 C:Genetics:

A:Gene: ATSP:F1707.15
 A:Map position: 1
 A:Introns: 118/3; 292/3; 373/3; 457/2; 826/3; 875/3; 902/3; 1057/3; 1093/2
 C:Superfamily: ubiquitin-protein ligase homology
 F:756-1120/Domain: ubiquitin-protein ligase homology <UBI>

Query Match 20.1%; Score 780.5; DB 2; Length 1126;
 Best Local Similarity 40.1%; Pred. No. 2.4e-48;
 Matches 170; Conservative 67; Mismatches 138; Indels 49; Gaps 6;
 A:Map position: 1

QY 343 DLVQKLVLRHLSLQPPQ--AGHCRIEVSREEIFEESYQIMKMRPKDLKKRLMVKFRG 400
 DB 707 DFDNKKAYFRSRIRHQDHISGRLISVRRAYVLEDSYNQLRMRSPODLKGRNLNVQFQG 766
 QY 401 BEGLDYGGVAREWMLYLICHEMLNPPYGLFOYSTDNIMYLQINPDSSINPDHLSYFHFVGR 460
 DB 767 BEGDAGLTREWYQLLSRVIFDKGALLFT-TVGNDAFQFNPNNSVYQTEHLSYFKFVGR 825
 QY 461 IMGLAVFHGHYINGGFTVPFYKQLLKPQLSDLESVDPELHKLWILENDITPVLDT 520
 DB 826 MVAKALPDQLLDVYFTRSPYKHLGVKTVHDIEAVDPDYKXKLWLENDVSDILDT 885
 QY 521 FCVEHNAFGRIL-----QHLEKPNGRNVPVTEENKKEYVLYVNWFRMRGIEAQLAL 573
 DB 886 FMSDADEKHILYKTEVTDYELKPGGRNIRVTEETHEVYVDLVADHILTSAIRPQINAF 945
 QY 574 QKGFNELIPQHLKPPDKOKELELIIGGLDKID----- 605
 DB 946 BEGLNELIPRELVSINDKELELLISGLPEIDCKLAFYSFLHLFAYSFKVITILLSV 1005
 QY 606 -----LNDWKNTRLKHCVADSNIVRWFQAVETFDERRARLLQFVTGSTRVPLQGFKA 660
 DB 1006 PFCELVDDLKANTETSYTAGSPVIRWFWEVVKAFSKEDMARFLQFVTGSTRVPLQGFKA 1065
 QY 661 LOGSTGAAGRLFTHLIDANTDLNPKAHTCFNIDIPPYESYEKLYEKLTAVEET--- 717
 DB 1066 LQ---GISGPORLQHKAYGAPERLPSAHTCFNQLDLPFYQSKEQLLTAHEANEG 1122
 QY 718 CGFA 721
 DB 1123 FGFA 1126

RESULT 14
 S69625
 hypothetical protein YDR457w - yeast (Saccharomyces cerevisiae)
 C:Species: Saccharomyces cerevisiae
 C:Date: 22-Aug-1995 #sequence_revision 06-Sep-1996 #text_change 23-Mar-2001
 A:Accession: S69625
 R:Dietrich, F.S.
 submitted to the EMBL Data Library, August 1995
 A:Description: The sequence of S. cerevisiae cosmid 9410, 8035, 8166, and 9787.
 A:Reference number: S69554
 A:Cross-references: EMBL:U3050; NID:927726; PIDN:AAB64910.1; PID:927738; MIPS:YDR457w
 A:Molecule type: DNA
 A:Residues: 1-3268 <DIE>
 A:Cross-references: EMBL:U3050; NID:927726; PIDN:AAB64910.1; PID:927738; MIPS:YDR457w
 C:Genetics:
 A:Gene: SGD:TOM1
 A:Cross-references: SGD:S0002865; MIPS:YDR457w
 A:Map position: 4R

Query Match 19.8%; Score 769; DB 2; Length 3268;
 Best Local Similarity 43.3%; Pred. No. 8.7e-47;
 Matches 152; Conservative 70; Mismatches 125; Indels 4; Gaps 2;
 A:Map position: 4R

QY 367 EVSREEIFEESYQIMKMRPKDLK-KELMVKFRGEGLDYGGVAREWMLYLICHEMLNPPY 425
 DB 2913 ITVREQVPLDSYALFFKTNDKSKLEITFKGESGVADAGVTRWEYQVLSRQMFND 2972
 QY 426 YGLFOYSTDNIMYLQINPDSSINPDHLSYFHFVGRIMGLAVFHGHYINGGTVFVKQLL 485
 DB 2973 YALFLPVSDEKTTTHPNRTSGINPEHLSFFKFIQMGKAIKRDQCFLDCHESREYKML 3032

```
QY 486 GHPISQSDLESVDPBLHSLVWILENDITPVLDHTFCVEHNAFGRILQHELKPNGRNVPV 545
Db 3033 GRPVSQKDMESLDPYKSLVWILENDITDIETTFVETDYGSHKVINIEGKDIIV 3092
QY 546 TRENKKEYVRLVYVNRKRGIEAQFLALQKGFNELIPQHLKPFQKELIIGGLDKID 605
Db 3093 TEANKQDYVKKVVEKYLQTSVKEQMDNPLVGFYALISKDLITIPDEQELELLISGLPDID 3152
QY 606 LNDWNSNTRLKHCVAADSNIVRWFQAVETFEERRARLLQFVTGSTRVPLQGFKALQGST 665
Db 3153 VDDWKNNTYVNTYTATCKEVSFWRAVSFDEAERAKLLQFVTGTSKVPPLNGFKELSGVN 3212
QY 666 GAAGRPLFTIHLIDANTONLPAKHCENRIDIPIPVSEYKLYEKLTLTAVEE 716
Db 3213 GVC---KFSIHRDFGSSERLPSSHICFNQNLNLPPIESYETLRGSLLLAINE 3260

RESULT 15
T20274
hypotheical protein F45H7.6 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 18-Aug-2000
C:Accession: T20274; T22257
R:Percy, C.
submitted to the EMBL Data Library, November 1994
A:Reference number: Z19245
A:Accession: T20274
A>Status: preliminary; translated from GB/EMBL/DBDJ
A:Molecule type: DNA
A:Residues: 1-889 <W1>
A:Cross-references: EMBL:Z46793; PIDN:CAA86773.1; GSPDB:GN00021; CESP:F45H7.6
A:Experimental source: clone C56G7
R:Percy, C.
submitted to the EMBL Data Library, June 1994
A:Reference number: Z19538
A:Accession: T22257
A>Status: preliminary; translated from GB/EMBL/DBDJ
A:Molecule type: DNA
A:Residues: 1-889 <W2>
A:Cross-references: EMBL:Z34800; PIDN:CAA84325.1; GSPDB:GN00021; CESP:F45H7.6
A:Experimental source: clone F45H7
C:Genetics:
A:Gene: CESP:F45H7.6
A:Map position: 3
A:Introns: 30/3; 54/1; 118/1; 185/2; 213/1; 260/3; 288/2; 353/2; 454/1; 559/2; 642/2; 74
C:Superfamily: WW repeat homology
F:229-266/Domain: WW repeat homology <W1>
F:372-408/Domain: WW repeat homology <W2>

Query Match 17.2%; Score 669; DB 2; Length 889;
Best Local Similarity 28.7%; Pred. No. 2.3e-40;
Matches 168; Conservative 104; Mismatches 203; Indels 110; Gaps 15;

QY 232 YECRTTVQGVYFLHTQTGVSTWHDPRIPRDLN-----SVNCDLGLPLPP 276
Db 322 YENQDAMQ-----IYNESVVRHAIHQKLDPPSKFENQPLFVFVNLFPADITQPLPS 376
QY 277 GWEVRSTVSGRIYFVDHNNRTTQTPRLHHMHQCLKPSQPLPLPSEGS----- 329
Db 377 GWEC-ITMNNRTVFLNHANKETSFYDPRIRF-----ETKTSRRGRSVFSRSSTAHKGI 430
QY 330 -----LEDEELP--AQRVERDLV----- 345
Db 431 DHALISKCDLRKIAODNFPQIAERISKMLLIERFGLAVASLANDLITLALSMLDSN 490
QY 346 -----QKLKVLRLHLSLQFPQAGHCRI--EVSREEIFESYQIMKMRPKDLKK-RL 394
Db 491 TEXLAGEGDNIKMFYEDMKKEKLGKPSRLCWKVSRRDLLDDAFRIILNVDPFVLKKSRL 550
QY 395 MVKPRGEEGLDYGVAREWLYLICHEMLNPYGLFQYSTNTIYMLQINPDSSINPDHLSY 454
Db 551 HIRFEGELALDYGGLSREFFILLRSLFHPKNGYFEY-NDYHLQLRPRGCETEKEKKW 609
```

```
QY 455 FHFVGRIMGLAVFHGYINGGFTVPFYKQLLGKPIQLSDLESVDPBLHSLVWILENDIT 514
Db 610 LILCGRVIALAVIHRCTIIVFFTNVYKSLQKRPVTLMDFKESDADFYSKMWLLENDVV 669
QY 515 PVLDTFCVEHNAFGRI-----LQHELKPNGRNVPVTEENKKEYVRLVYN 559
Db 670 D-LEMSFVYSVMVNGKVSDDLTLSSIVSVALQAEQLVPGGESQMVTEANKAEFIDLMCQ 728
QY 560 WRPMRGIEAQFLALQKGFNELIPQHLKPFQKELIIGGLDKIDLNDWKSNTRLKHCV 619
Db 729 KKAIRGVKEKPLEILLTSFNQILNDNLNSLESSDLKRLISGSELDLNDWRTWTIYKGY 788
QY 620 ADSNI-VRFWFQAVETFEERRARLLQFVTGSTRVPLQGFKALQGSTGAAGPLFTIHLI 678
Db 789 SDCHIVVWEVWEVETMTNQERFDLLLFVTGSSVVPFEGFSALRGNEIS---KFCIEKW 845
QY 679 DANTDNLPAKHTCFENRIDIPIPVSEYKLYEKLTLTAVEETCGFAVE 723
Db 846 -GDATSPRAHTCFNRLQLPSYNTKQQLKSKLOQAIYVNGMSYSIE 889

Search completed: September 21, 2004, 07:48:25
Job time : 35.9986 secs
```

This Page Blank (uspio)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 21, 2004, 07:29:24 : Search time 16.7111 Seconds
(without alignments)
2252.800 Million cell updates/sec

Title: US-10-009-945-2

Perfect score: 3884

Sequence: 1 GSSSKIRLTVCANLAKK.....EKLYEKLITAVEETCGFAVE 723

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_42:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	3855	99.3	757	1	SUFL_HUMAN	O9hce7 homo sapien
2	3548	91.3	731	1	SUFL_XENLA	O9pun2 xenopus lae
3	3249.5	83.7	619	1	SUFL_MOUSE	O9cun6 mus musculus
4	3001	77.3	748	1	SUFL_HUMAN	O9hau4 homo sapien
5	1492.5	38.4	767	1	PUB1_SCHPO	O92462 schizosacch
6	1377	35.5	809	1	RSP5_YEAST	P39940 saccharomyc
7	1318	33.9	864	1	ITCH_MOUSE	O8c863 mus musculus
8	1288.5	33.2	903	1	ITCH_HUMAN	O96102 homo sapien
9	1259	32.4	1000	1	NER4_HUMAN	P46934 homo sapien
10	1239	31.9	922	1	WP21_HUMAN	O9homo homo sapien
11	1234	31.8	870	1	WP21_MOUSE	O9dbh0 mus musculus
12	1232	31.7	916	1	WP21_MOUSE	O8bzz3 mus musculus
13	1227.5	31.6	887	1	NER4_RAT	Q62940 rattus norv
14	1225	31.5	870	1	WP21_HUMAN	O00308 homo sapien
15	1225	31.5	887	1	NER4_MOUSE	P46935 mus musculus
16	627.5	16.2	210	1	URB1_RAT	P51593 rattus norv
17	534	13.7	875	1	UE3A_HUMAN	Q05086 homo sapien
18	483	12.4	885	1	UE3A_MOUSE	O08759 mus musculus
19	481.5	12.4	892	1	HUL4_YEAST	P40985 saccharomyc
20	449.5	11.6	1050	1	HER3_HUMAN	Q15034 homo sapien
21	358	9.2	1992	1	TRB_HUMAN	Q14669 homo sapien
22	354	9.1	910	1	HUL5_YEAST	P53119 saccharomyc
23	278.5	7.2	2799	1	EDD_HUMAN	O95071 homo sapien
24	277	7.1	920	1	EDD_RAT	Q64671 rattus norv
25	262.5	6.8	1483	1	UF04_YEAST	P32202 saccharomyc
26	257	6.6	1647	1	YDE1_SCHPO	Q10435 schizosacch
27	238.5	6.1	472	1	YAP1_MOUSE	P46938 mus musculus
28	238.5	6.1	1620	1	HED1_HUMAN	O9ult8 homo sapien
29	222	5.7	2895	1	HYD_DROME	P51592 drosophila
30	173.5	4.5	1277	1	AIF1_RAT	O88382 rattus norv
31	172.5	4.4	1455	1	AIF1_HUMAN	Q86u18 homo sapien
32	172	4.4	1275	1	AIF1_MOUSE	Q9wvq1 mus musculus
33	162	4.2	448	1	YAP1_CHICK	P46936 gallus gall

ALIGNMENTS

RESULT 1

```

SUFL_HUMAN          STANDARD;          PRT;   757 AA.
ID   SUFL_HUMAN          454      1   YAP1_HUMAN
34  161.5      4.2      454      1   YAP1_HUMAN
35  143.5      3.7      383      1   SAVI_HUMAN
36  143.5      3.7      386      1   SAVI_MOUSE
37  128.5      3.3      1555     1   RIM2_RAT
38  128      3.3      1813     1   UN13_CABEL
39  119.5      3.1      1530     1   RIM2_MOUSE
40  119      3.1      1575     1   BAG3_HUMAN
41  118      3.0      1188     1   RIM2_HUMAN
42  116.5      3.0      577     1   BAG3_MOUSE
43  115.5      3.0      1692     1   RIM1_HUMAN
44  114.5      2.9      803     1   RSG5_HUMAN
45  114      2.9      593     1   CNE5_HUMAN

P46937 homo sapien
Q9h4b6 homo sapien
Q8veb2 mus musculus
Q9j1s1 rattus norv
P27715 caenorhabdi
Q9eg7 mus musculus
Q95817 homo sapien
Q9ug26 homo sapien
Q9j1v1 mus musculus
Q86ur5 homo sapien
O43374 homo sapien
Q9hch3 homo sapien

ALIGNMENTS

SEQUENCE OF 10-731 FROM N.A. (ISOFORM SHORT).
MEDLINE=99385348; PubMed=10459166;
Zhu H., Kavaak P., Abdollah S., Wrana J.L., Thomsen G.H.;
RT "A SMAD ubiquitin ligase targets the BMP pathway and affects embryonic
pattern formation.";
RL Nature 400:687-693(1999).
[2]
SEQUENCE OF 20-731 FROM N.A. (ISOFORM SHORT).
Stoneking T., Bauer C., O'Neal D.;
Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.
[3]
SEQUENCE FROM N.A. (ISOFORM LONG).
TISSUE=Brain;
MEDLINE=20450683; PubMed=10997877;
Nagase T., Kikuno R., Nakayama M., Hirosewa M., Ohara O.;
RT "Prediction of the coding sequences of unidentified human genes.
XVIII. The complete sequences of 100 new cDNA clones from brain which
code for large proteins in vitro.";
DNA Res. 7:273-281(2000).
CC -!- FUNCTION: Interacts with receptor-regulated SMADs specific for the
bmp pathway (SMAD1 AND SMAD5) in order to trigger their
ubiquitination and degradation and hence their inactivation.
CC -!- ALTERNATIVE PRODUCTS:
Event=Alternative splicing; Named isoforms=2;
Name=Long;
IsoId=Q9HCE7-1; Sequence=Displayed;
Name=Short;
IsoId=Q9HCE7-2; Sequence=VSP_006812;
CC -!- SIMILARITY: Contains 1 C2 domain.
CC -!- SIMILARITY: Contains 2 WW domains.
CC -!- SIMILARITY: Contains 1 HECT-type E3 ubiquitin-protein ligase
domain.
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).

```

```

CC -----
DR EMBL: AF199364; AAF08298.2; -
DR EMBL: AC004893; AAC62434.1; -
DR EMBL: AB046845; BAB13451.1; ALT_INIT.
DR HSPF: Q13526; 1FN.
DR MIM: 605568; -
DR GO: GO:0005622; C:intracellular; TAS.
DR GO: GO:0000211; F:protein degradation tagging activity; IDA.
DR GO: GO:0004842; F:ubiquitin-protein ligase activity; IDA.
DR GO: GO:0030154; P:cell differentiation; TAS.
DR GO: GO:0007398; P:ectoderm development; IDA.
DR GO: GO:0030314; P:negative regulation of BMP signaling pathway; TAS.
DR GO: GO:0006512; P:ubiquitin cycle; IDA.
DR InterPro: IPR000008; C2.
DR InterPro: IPR008973; C2 CalB.
DR InterPro: IPR000569; HECT domain.
DR InterPro: IPR001202; WW_Rsp5_WWP.
DR Pfam: PF00168; C2_1.
DR Pfam: PF06632; HECT_1.
DR Pfam: PF00397; WW_2.
DR SMART: SM00239; C2; 1.
DR SMART: SM00119; HECTC; 1.
DR SMART: SM00456; WW_2.
DR PROSITE: PS00499; C2 DOMAIN 1; 1.
DR PROSITE: PS00004; C2 DOMAIN_2; 1.
DR PROSITE: PS0237; HECT; 1.
DR PROSITE: PS01159; WW DOMAIN 1; 1.
DR PROSITE: PS00020; WW DOMAIN_2; 2.
DR KEGG: Ubl conjugation pathway; Ligase; Repeat; Alternative splicing.
FW DOMAIN 1 99 C2 DOMAIN.
FT DOMAIN 234 267 WW 1.
FT DOMAIN 306 339 WW 2.
FT BINDING 420 757 HECT.
FT BINDING 725 725 UBIQUITIN.
FT VARSPIC 269 294 Missing (in isoform Short).
FT MUTAGEN 725 725 /FTID=VSP_006912.
FT SEQUENCE 757 AA; 86113 MW; 89A17CFC47B40E9 CRC64;
Query Match 99.3%; Score 3855; DB 1; Length 757;
Best Local Similarity 96.5%; Pred. No. 3.2e-274;
Matches 722; Conservative 0; Mismatches 0; Indels 26; Gaps 1;
QY 2 GSSIKIRLVLCANLAKKQFFRLDPDFAKIVVDGSGQCHSTDTVKNTLDPKKNQHYDLY 61
DB 10 GSSIKIRLVLCANLAKKQFFRLDPDFAKIVVDGSGQCHSTDTVKNTLDPKKNQHYDLY 69
QY 62 VGKTDSTISVNNHKKIKKQAGFLGCVRLLSNAISRLKDTGYQRLDLCKLNPSTDAV 121
DB 70 VGKTDSTISVNNHKKIKKQAGFLGCVRLLSNAISRLKDTGYQRLDLCKLNPSTDAV 129
QY 122 RGQIVLSLQTRDRTGTGSGVDDCGLLNKGTVYEDSGGRPLSCFEEBPAYTDSGAA 181
DB 130 RGQIVLSLQTRDRTGTGSGVDDCGLLNKGTVYEDSGGRPLSCFEEBPAYTDSGAA 189
QY 182 AGGNCNRFVSPSQDLQRLNPDVRSGLQTPQNRPHGHOSPELPEGYEQRTTVQGG 241
DB 190 AGGNCNRFVSPSQDLQRLNPDVRSGLQTPQNRPHGHOSPELPEGYEQRTTVQGG 249
QY 242 VYFLHTGTGVTWHDRI - - - - - PRDLNSVNCDELGLPL 275
DB 250 VYFLHTGTGVTWHDRI - - - - - PRDLNSVNCDELGLPL 309
QY 276 PGWEVRSTVSGRIYFVDHNNRTTQTPRLHHNNHQCQLKPSQPLPSEGSLEDEEL 335
DB 310 PGWEVRSTVSGRIYFVDHNNRTTQTPRLHHNNHQCQLKPSQPLPSEGSLEDEEL 369
QY 336 PAQYERDLVQKLVLRHLSLQOPQACHRIEVSREIFEESYRQIMKRPDLKKRLM 395
DB 370 PAQYERDLVQKLVLRHLSLQOPQACHRIEVSREIFEESYRQIMKRPDLKKRLM 429
QY 396 VKFRGEGLDYGVAEWLYLLCHEMLNPNYGLFQYSTDNIMYMLQINPDSSINPDHLSYF 455

```

```

DB 430 VKFRGEGLDYGVAEWLYLLCHEMLNPNYGLFQYSTDNIMYMLQINPDSSINPDHLSYF 489
QY 456 HPVGRIMGLAVFHGHYINGGFTVPFYKOLLGKPIQLSDLESVDPELHKSLVWILENDITP 515
DB 490 HPVGRIMGLAVFHGHYINGGFTVPFYKOLLGKPIQLSDLESVDPELHKSLVWILENDITP 549
QY 516 VLDHTFCVHNAPGRILQHELKPNGRNVPVTEENKKEYVRLYVNWRFMRGTEAFLALQK 575
DB 550 VLDHTFCVHNAPGRILQHELKPNGRNVPVTEENKKEYVRLYVNWRFMRGTEAFLALQK 609
QY 576 GNEELIPOHLKPPDQKELELIIGLDKIDLNKSNTRLKHCVCADSNIVFWFQAVET 635
DB 610 GNEELIPOHLKPPDQKELELIIGLDKIDLNKSNTRLKHCVCADSNIVFWFQAVET 669
QY 636 DEERARLLQFVTGSTRVPLQGFKALQGSTGAAGPRLFTIHLIDANTDNLKHAHTCFNRI 695
DB 670 DEERARLLQFVTGSTRVPLQGFKALQGSTGAAGPRLFTIHLIDANTDNLKHAHTCFNRI 729
QY 696 DIPPYESYEKLYEKLITAVEETCGFAVE 723
DB 730 DIPPYESYEKLYEKLITAVEETCGFAVE 757

RESULT 2
SUP1_XENLA
ID SUP1_XENLA STANDARD; PRT; 731 AA.
AC Q9PUN2;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Smad ubiquitination regulatory factor 1 (EC 6.3.2.-) (Ubiquitin--
GN protein ligase SMURF1) (Smad-specific E3 ubiquitin ligase) (xSMURF1).
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidea; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]_TaxID=8355;
RP SEQUENCE FROM N.A.
RC TISSUE=Blascula;
RX MEDLINE=99385348; PubMed=10458166;
RT Zhu H., Kavsak P., Abdollah S., Wrana J.L., Thomsen G.H.;
RT "A SMAD ubiquitin ligase targets the BMP pathway and affects embryonic
RT pattern formation.";
RL Nature 400:687-693 (1999).
CC -!- FUNCTION: INTERACTS WITH RECEPTOR-REGULATED SMADS SPECIFIC FOR THE
CC BMP PATHWAY (SMAD1 AND SMAD5) IN ORDER TO TRIGGER THEIR
CC UBIQUITINATION AND DEGRADATION AND HENCE THEIR INACTIVATION. MAY
CC REGULATE ECTODERMAL DIFFERENTIATION AND PATTERN BY MODULATING BMP
CC SIGNALING AND MAY ENHANCE CELLULAR RESPONSIVENESS TO THE SMAD2
CC (ACTIVIN/TGF-BETA) PATHWAY.
CC -!- DEVELOPMENTAL STAGE: EXPRESSED FROM THE EGG STAGE TO THE SWIMMING
CC TADPOLE, WITH MAXIMUM LEVELS OBSERVED IN THE STAGES FROM EGG TO
CC GASTRULA. AT GASTRULATION DISTRIBUTED UNIFORMLY IN EMBRYONIC
CC ECTODERM AND INVOLUTING MESODERM, AND EXPRESSION GRADUALLY
CC LOCALIZES TO THE NERVOUS SYSTEM. AT EARLY TADPOLE STAGES EXPRESSED
CC IN THE CNS, EYE, BRANCHIAL ARCHES, KIDNEY AND SOMITES.
CC -!- SIMILARITY: Contains 1 C2 domain.
CC -!- SIMILARITY: Contains 2 WW domains.
CC -!- SIMILARITY: Contains 1 HECT-type E3 ubiquitin-protein ligase
CC domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.ebi.ac.uk/ebis/
CC or send an email to license@ebi.ac.uk).
CC -----
CC EMBL; AF169310; AAD52564.1; -
CC GO; GO:0005622; C:intracellular; TAS.

```

[illegible]

modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

 CC EMBL; AK015264; BAB29770.2; ALT_INIT.
 CC HSP; Q13526; 1PIN.
 CC MGD; MGI11923038; 4930431E10Rik.
 CC GO; GO:0005622; C:intracellular; ISS.
 CC GO; GO:0002111; F:protein degradation tagging activity; ISS.
 CC GO; GO:0004842; F:ubiquitin-protein ligase activity; ISS.
 CC GO; GO:0030154; P:cell differentiation; ISS.
 CC GO; GO:0007398; P:ectoderm development; ISS.
 CC GO; GO:0030514; P:negative regulation of BMP signaling pathway; ISS.
 CC GO; GO:0006512; P:ubiquitin cycle; ISS.
 CC InterPro; IPR000008; C2.
 CC InterPro; IPR000569; HECT domain.
 CC InterPro; IPR001202; WW_Rsp5_WMP.
 CC Pfam; PF00632; HECT; 1.
 CC Pfam; PF00397; WW; 2.
 CC SMART; SM00119; HECTC; 1.
 CC SMART; SM00456; WW; 2.
 CC PROSITE; PS00499; C2_DOMAIN_1; PARTIAL.
 CC PROSITE; PS00004; C2_DOMAIN_2; PARTIAL.
 CC PROSITE; PS0237; HECT; 1.
 CC PROSITE; PS01159; WW_DOMAIN_1; 1.
 CC PROSITE; PS00200; WW_DOMAIN_2; 2.
 CC Ubl conjugation pathway; Ligase; Repeat.
 KW NON TER 1
 FT DOMAIN 125 158 WW 1.
 FT DOMAIN 171 204 WW 2.
 FT DOMAIN 285 619 HECT.
 FT DOMAIN 35 38 POLY-GLY.
 SQ SEQUENCE 619 AA; 70902 MW; 4CBE2F8624A7B525 CRC64;

Query Match 83.7%; Score 3249.5; DB 1; Length 619;
 Best Local Similarity 97.3%; Pred. No. 4.9e-230;
 Matches 605; Conservative 6; Mismatches 8; Indels 3; Gaps 1;

QY 102 DTGQRLDLCKLNPSDITDAVREGQVIVSLQTRDRIGTGGVSDRCGLNENGTVYDSGPG 161
 DB 1 DTGQRLDLCKLNPSDITDAVREGQVIVSLQTRDRIGTGGVSDRCGLNENGTVYDSGPG 60
 QY 162 RPLSCFEEPAFYDSTGAAAGGNCRFVSPSQRLQRLQRLNPVRSIQTQNRPH 221
 DB 61 RPLSCLMEEPAYDGTGAAAGGNCRFVSPSQRLVQLQRLNPVRSIQTQNRPH 120
 QY 222 GHSQPELPEGEQRTTQGGQVYFHTGTGVTWHDPRIPDLNSVNCDELGLPPLPGWEVR 281
 DB 121 GHSQPELPEGEQRTTQGGQVYFHTGTGVTWHDPRIPDLNSVNCDELGLPPLPGWEVR 180
 QY 282 STVSGRIYFVDHNRRTQFTDPRLLHMHQCQLKEPSQPLPSPESGLEDDELPAQRYE 341
 DB 181 STVSGRIYFVDHNRRTQFTDPRLLHMHQCQLKEPSQPLPSPESGLEDDELPAQRYE 240
 QY 342 RDLVQKLKVLREHLSLQOPQAGHCRIEVSREEIPEESYVIMKMRPDLAKRLMVKFRGE 401
 DB 241 RDLVQKLKVLREHLSLQOPQAGHCRIEVSREEIPEESYVIMKMRPDLAKRLMVKFRGE 300
 QY 402 EGLDYGGVAREWLXLLCHEMLNPYYGLFQYSTDNIMYLQINPDSSINPDHLSYFHFVGR 461
 DB 301 EGLDYGGVAREWLXLLCHEMLNPYYGLFQYSTDNIMYLQINPDSSINPDHLSYFHFVGR 360
 QY 462 MGLAVFHGHYNGGFTVPFYKQLGKPIQLSDLESVDPELHKSILVILENDITPVLDTHT 521
 DB 361 MGLAVFHGHYNGGFTVPFYKQLGKPIQLSDLESVDPELHKSILVILENDITPVLDTHT 420
 QY 522 CVEHNAFRIILCHELKPGRNPVPTENKKEYVRLYVNRFMRGIEAQFLALQKGFNEL 581
 DB 421 CVEHNAFRIILCHELKPGRNPVPTENKKEYVRLYVNRFMRGIEAQFLALQKGFNEL 480
 QY 582 PQHLLKPPDQKELELIIGGLDKIDLNDKSNTRLKHCVDASNIVRWQAVTFDEERRA 641
 DB 481 PQHLLKPPDQKELELIIGGLDKIDLNDKSNTRLKHCVDASNIVRWQAVTFDEERRA 540

QY 642 RLLOFVTGSTRVLOGKALOGSTGAAGPRLFTIHLIDANTDNLKPAHTCNFRIDIPPE 701
 DB 541 RLLOFVTGSTRVLOGKALOGSTGAAGPRLFTIHLIDANTDNLKPAHTCNFRIDIPPE 597
 QY 702 SYEKLYEKLTAVEETCGFAVE 723
 DB 598 SYEKLYEKLTAVEETCGFAVE 619
 RESULT 4
 SUP2 HUMAN
 ID SUP2 HUMAN STANDARD; PRT; 748 AA.
 AC Q9HAU4; Q9H260;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Smad ubiquitination regulatory factor 2 (EC 6.3.2.-) (Ubiquitin--
 DE protein ligase SMURF2) (Smad-specific E3 ubiquitin ligase 2)
 DE (hSMURF2)
 GN SMURF2
 OS Homo sapiens (Human)
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A., AND MUTAGENESIS OF 251-PRO--VAL-284 AND
 RP 297-GLY--LEU-330.
 RX PubMed=1163210;
 RA Kavsak P., Rasmussen R.K., Causing C.G., Bonni S., Zhu H.,
 RA Thomsen G.H., Wrana J.L.;
 RT "Smad7 binds to Smurf2 to form an E3 ubiquitin ligase that targets the
 RT TGF-beta receptor for degradation.";
 RL Mol. Cell 6:1365-1375(2000).
 RN [2]
 RP SEQUENCE FROM N.A., AND MUTAGENESIS OF 251-PRO--VAL-284 AND CYS-716.
 RP MEDLINE=20538422; PubMed=11015919;
 RA Lin X., Liang M., Feng X.-H.;
 RT "Smurf2 is a ubiquitin E3 ligase mediating proteasome-dependent
 RT degradation of Smad2 in transforming growth factor-beta signaling.";
 RL J. Biol. Chem. 275:36818-36822(2000).
 RN [3]
 RP SEQUENCE FROM N.A., AND MUTAGENESIS OF CYS-716.
 RP MEDLINE=21107656; PubMed=11158580;
 RA Zhang Y., Chang C., Gehring D.J., Hemmati-Briylanlou A., Derynck R.;
 RT "Regulation of Smad degradation and activity by Smurf2, an E3
 RT ubiquitin ligase.";
 RL Proc. Natl. Acad. Sci. U.S.A. 98:974-979(2001).
 CC -!- FUNCTION: Interacts with SMAD1, SMAD2 and SMAD7 in order to
 CC trigger their ubiquitination and proteasome-dependent degradation.
 CC Enhances the inhibitory activity of SMAD7 and reduces the
 CC transcriptional activity of SMAD2. Coexpression of SMURF2 with
 CC SMAD1 results in considerable decrease in steady-state level of
 CC SMAD1 protein and a smaller decrease of SMAD2 level.
 CC -!- SUBUNIT: Interacts with SMAD1, SMAD2, SMAD3, SMAD6 and SMAD7 but
 CC not SMAD4.
 CC -!- SUBCELLULAR LOCATION: Nuclear. Cytoplasmic in the presence of
 CC SMAD7.
 CC -!- TISSUE SPECIFICITY: Widely expressed.
 CC -!- DOMAIN: The second and third WW domains are responsible for
 CC interaction with R-SMAD (SMAD1, SMAD2 and SMAD3).
 CC -!- SIMILARITY: Contains 1 C2 domain.
 CC -!- SIMILARITY: Contains 3 WW domains.
 CC -!- SIMILARITY: Contains 1 HECT-type E3 ubiquitin-protein ligase
 CC domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

```

CC -----
DR EMBL: AF310676; AAG45422.1; --
DR EMBL: AF301463; AAG25641.1; --
DR EMBL: AY014180; AAG50421.1; --
DR HSP: Q13526; 1PIN.
DR MIM: 605532; --
DR GO: GO:0004842; F-ubiquitin-protein ligase activity; NAS.
DR GO: GO:0016481; P-negative regulation of transcription; NAS.
DR GO: GO:0017015; P-regulation of TGFbeta receptor signaling pa. .; NAS.
DR GO: GO:0006511; P-ubiquitin-dependent protein catabolism; NAS.
DR InterPro: IPR000008; C2.
DR InterPro: IPR008973; C2_CalB.
DR InterPro: IPR000569; HECT domain.
DR InterPro: IPR001202; WW_Rsp_WWP.
DR Pfam: PF00168; C2; 1.
DR Pfam: PF00632; HECT; 1.
DR Pfam: PF00397; WW; 3.
DR SMART: SM00239; C2; 1.
DR SMART: SM00119; HECTC; 1.
DR SMART: SM00456; WW; 3.
DR PROSITE: PS00499; C2_DOMAIN_1; 1.
DR PROSITE: PS00004; C2_DOMAIN_2; 1.
DR PROSITE: PS02337; HECT; 1.
DR PROSITE: PS01159; WW_DOMAIN_1; 1.
DR PROSITE: PS00020; WW_DOMAIN_2; 3.
KW Ub1 conjugation pathway; Ligase; Repeat; Nuclear protein.
FT DOMAIN 1 98
FT DOMAIN 157 190 WW 1.
FT DOMAIN 251 284 WW 2.
FT DOMAIN 297 330 WW 3.
FT DOMAIN 414 748 HECT.
FT MUTAGEN 251 284 MISSING: ABOLISHES INTERACTION WITH
FT MUTAGEN 297 330 SWAD2 AND SWAD7.
FT MUTAGEN 716 716 MISSING: ABOLISHES INTERACTION WITH
FT MUTAGEN 716 716 C->A: LOSS OF ABILITY TO UBIQUITINATE
FT MUTAGEN 716 716 C->G: LOSS OF ACTIVITY (LOSS OF ABILITY
FT MUTAGEN 716 716 TO UBIQUITINATE SWAD1 AND SWAD2 AND NO
FT MUTAGEN 716 716 DOWN-REGULATION OF SWAD1 AND SWAD2
FT MUTAGEN 716 716 PROTEIN LEVELS).
FT CONFLICT 6 6 G -> R (IN REF. 2).
FT SEQUENCE 748 AA; 86195 NW; 3042B443A3755762 CRC64;
Query Match 77.3%; Score 3001; DB 1; Length 748;
Best Local Similarity 74.6%; Pred. No. 1e-211;
Matches 565; Conservative 63; Mismatches 71; Indels 58; Gaps 8;
QY 5 IKRLTVLCAKNAKDKFFRLPDPFAKIVVDGSGQCHSTDTVXNTLDPKWNQHYDLYVGK 64
DB 12 VKRLTVLCAKNAKDKFFRLPDPFAKIVVDGSGQCHSTDTVXNTLDPKWNQHYDLYGK 71
QY 65 TDSITISVNNHKIKKQAGLGLCVRLLSNAISLKTGTQRLDCKLNPSDQDAVRGQ 124
DB 72 SDSVITISVNNHKIKKQAGLGLCVRLLSNAISLKTGTQRLDCKLGNPDNDTVRGQ 131
QY 125 IVVSLQTRDRIQTGGVSDCRGLLENE-----GTVY----- 155
DB 132 IVVSLQSRDRIQTGGVSDCRGLFNDLDPDQWERTASGLIYLNHITRTQWERTP 191
QY 156 --EDSGRPLSCFWEPEAPYTDSTGAAGGNCRFVESPQDQRLQRLNPDVRSGL 213
DB 192 ASEYSGRPLSCFVDENTPISGTNGATCG-----QSSDPLAERRVRSQRHNYM 242
QY 214 QTPQNRPHQHQSPQLPEGVQRSTTVQGVYFLHTGTGVSTWHDDRIPELNSVNCDELGP 273
DB 243 ----SRTHLHTPPDPEGVQRSTTVQGVYFLHTGTGVSTWHDDRIPELNSVNCDELGP 298
QY 274 LPPGHEVSTVSGRIYFVDHNNRTTQFTDPR-----LHHMNHQCOLKEPSPQLPSEGS 329
DB 299 LPPGWEIRNTATGRVYFVDHNNRTTQFTDPRLSANLHLVLRNQLKQDQQQVV-----S 354
QY 330 L---EDELPAQRYERDLVQKLKVLRLHLSLQDQPAQAGHCRIEVSRREIFESYQIMKVR 386

```

```

DB 355 LCPDDTECLAVPRYKRLDVQKLKILRQELSQQPOAGHCRIEVSRREIFESYQIMKVR 414
QY 387 PKDLKRLMKVPRGEGLDYGGVAREWLYLLCHEMLNPPYGLFOYSTDNTYMLQINPDSS 446
DB 415 PKDLKRLMKVPRGEGLDYGGVAREWLYLLCHEMLNPPYGLFOYSRDDTYTLQINPDSA 474
QY 447 INPHLSYFFHVGIMGLAVFHGHYNGFTVFFYKQLIGKPIQLSDLESVDDELHKSIV 506
DB 475 VNPHELSYFFHVGIMGLAVFHGHYNGFTVFFYKQLIGKPIQLSDLESVDDELHKSIV 534
QY 507 WILENDITPVDHTFCVEHNAFGRILLOHELKPNRNVPTTEKNKYYRILYVNMFRMGI 566
DB 535 WILENDITGLVDHTFCVEHNAFGEIILQHELKPNKGIPTVNEENKYYRILYVNMFRMGI 594
QY 567 EAQFLALQKGFNELIPOHLKLPDOKLELIIGGLDKIDLNDKSNRLKHCVCVADSNI 626
DB 595 EAQFLALQKGFNEVPOHLKLPDOKLELIIGGLDKIDLNDKSNRLKHCVCVADSNI 654
QY 627 WFWQAVETFEERRARLLQFVTGSTRVPLQGFKALQSGTGAAGPRLFTIHLIDANTDNL 686
DB 655 WFWKAVEFFDEERRARLLQFVTGSTRVPLQGFKALQSGTGAAGPRLFTIHLIDANTDNL 711
QY 687 KAHTCFNRIDIPPEYSEKLYEKLTAIVETCGFAVE 723
DB 712 KAHTCFNRIDIPPEYSEKLYEKLTAIVETCGFAVE 748
RESULT 5
PUBL SCHPO
ID_PUBL SCHPO STANDARD; PRT; 767 AA.
AC Q92462; O14454;
DT 01-NOV-1997 (Rel. 35, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Ubiquitin-protein ligase pub1 (EC 6.3.2.-).
GN PUB1 OR SPAC11G7.02.
OS Schizosaccharomyces pombe (Fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
OX NCBI_TaxID=4896;
[1]
RN SEQUENCE FROM N.A.
RP MEDLINE=96205868; PubMed=8635463;
RA Nefsky B., Beach D.;
RT "Pub1 acts as an E3-AP-like protein ubiquitin ligase in the
RT degradation of cdc25."
RL EMBO J. 15:1301-1312(1996).
[2]
RN SEQUENCE FROM N.A.
RP STRAIN=J227;
RC MEDLINE=97340937; PubMed=9197411;
RA Saleki R., Jia Z., Karagiannis J., Young P.G.;
RT "Tolerance of low pH in Schizosaccharomycetes pombe requires a
RT functioning pub1 ubiquitin ligase."
RL Mol. Gen. Genet. 254:520-528(1997).
[3]
RN SEQUENCE FROM N.A.
RP STRAIN=972;
RC MEDLINE=21848401; PubMed=11859360;
RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
RA Sgouras K., Peat N., Hayles J., Baker S., Basham D., Bowman S.,
RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
RA Collins M., Connor R., Cronin A., Davis P., Feltwell T., Fraser A.,
RA Gentles S., Goble A., Hamlin N., Harris D., Hidaigo J., Hodgson G.,
RA Holroyd K., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagsis K.,
RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
RA Mooney P., Moutle S., Mungall K., Murphy L., Niblett D., Odell C.,
RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitsch E.,
RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,
RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,

```


DR	SGD; S0000327; RSP5.	DR	InterPro; IPR000008; C2.	DR	PRINTS; PRO0403; WWDOMAIN.	DR	SMART; SM00239; C2; 1.	DR	SMART; SM00119; HECTG; 1.	DR	SMART; SM00456; WW; 3.	DR	PROSITE; PS00499; C2_DOMAIN 1; 1.	DR	PROSITE; PS00004; C2_DOMAIN 2; 1.	DR	PROSITE; PS0237; HECT; 1.	DR	PROSITE; PS01159; WW_DOMAIN 1; 3.	DR	PROSITE; PS00200; WW_DOMAIN 2; 3.	KW	Ubl conjugation pathway; Ligase; Repeat.	FT	DOMAIN 1 88	FT	DOMAIN 2 88	FT	DOMAIN 3 88	FT	DOMAIN 4 88	FT	DOMAIN 5 88	FT	DOMAIN 6 88	FT	DOMAIN 7 88	FT	DOMAIN 8 88	FT	DOMAIN 9 88	FT	DOMAIN 10 88	FT	DOMAIN 11 88	FT	DOMAIN 12 88	FT	DOMAIN 13 88	FT	DOMAIN 14 88	FT	DOMAIN 15 88	FT	DOMAIN 16 88	FT	DOMAIN 17 88	FT	DOMAIN 18 88	FT	DOMAIN 19 88	FT	DOMAIN 20 88	FT	DOMAIN 21 88	FT	DOMAIN 22 88	FT	DOMAIN 23 88	FT	DOMAIN 24 88	FT	DOMAIN 25 88	FT	DOMAIN 26 88	FT	DOMAIN 27 88	FT	DOMAIN 28 88	FT	DOMAIN 29 88	FT	DOMAIN 30 88	FT	DOMAIN 31 88	FT	DOMAIN 32 88	FT	DOMAIN 33 88	FT	DOMAIN 34 88	FT	DOMAIN 35 88	FT	DOMAIN 36 88	FT	DOMAIN 37 88	FT	DOMAIN 38 88	FT	DOMAIN 39 88	FT	DOMAIN 40 88	FT	DOMAIN 41 88	FT	DOMAIN 42 88	FT	DOMAIN 43 88	FT	DOMAIN 44 88	FT	DOMAIN 45 88	FT	DOMAIN 46 88	FT	DOMAIN 47 88	FT	DOMAIN 48 88	FT	DOMAIN 49 88	FT	DOMAIN 50 88	FT	DOMAIN 51 88	FT	DOMAIN 52 88	FT	DOMAIN 53 88	FT	DOMAIN 54 88	FT	DOMAIN 55 88	FT	DOMAIN 56 88	FT	DOMAIN 57 88	FT	DOMAIN 58 88	FT	DOMAIN 59 88	FT	DOMAIN 60 88	FT	DOMAIN 61 88	FT	DOMAIN 62 88	FT	DOMAIN 63 88	FT	DOMAIN 64 88	FT	DOMAIN 65 88	FT	DOMAIN 66 88	FT	DOMAIN 67 88	FT	DOMAIN 68 88	FT	DOMAIN 69 88	FT	DOMAIN 70 88	FT	DOMAIN 71 88	FT	DOMAIN 72 88	FT	DOMAIN 73 88	FT	DOMAIN 74 88	FT	DOMAIN 75 88	FT	DOMAIN 76 88	FT	DOMAIN 77 88	FT	DOMAIN 78 88	FT	DOMAIN 79 88	FT	DOMAIN 80 88	FT	DOMAIN 81 88	FT	DOMAIN 82 88	FT	DOMAIN 83 88	FT	DOMAIN 84 88	FT	DOMAIN 85 88	FT	DOMAIN 86 88	FT	DOMAIN 87 88	FT	DOMAIN 88 88	FT	DOMAIN 89 88	FT	DOMAIN 90 88	FT	DOMAIN 91 88	FT	DOMAIN 92 88	FT	DOMAIN 93 88	FT	DOMAIN 94 88	FT	DOMAIN 95 88	FT	DOMAIN 96 88	FT	DOMAIN 97 88	FT	DOMAIN 98 88	FT	DOMAIN 99 88	FT	DOMAIN 100 88	FT	DOMAIN 101 88	FT	DOMAIN 102 88	FT	DOMAIN 103 88	FT	DOMAIN 104 88	FT	DOMAIN 105 88	FT	DOMAIN 106 88	FT	DOMAIN 107 88	FT	DOMAIN 108 88	FT	DOMAIN 109 88	FT	DOMAIN 110 88	FT	DOMAIN 111 88	FT	DOMAIN 112 88	FT	DOMAIN 113 88	FT	DOMAIN 114 88	FT	DOMAIN 115 88	FT	DOMAIN 116 88	FT	DOMAIN 117 88	FT	DOMAIN 118 88	FT	DOMAIN 119 88	FT	DOMAIN 120 88	FT	DOMAIN 121 88	FT	DOMAIN 122 88	FT	DOMAIN 123 88	FT	DOMAIN 124 88	FT	DOMAIN 125 88	FT	DOMAIN 126 88	FT	DOMAIN 127 88	FT	DOMAIN 128 88	FT	DOMAIN 129 88	FT	DOMAIN 130 88	FT	DOMAIN 131 88	FT	DOMAIN 132 88	FT	DOMAIN 133 88	FT	DOMAIN 134 88	FT	DOMAIN 135 88	FT	DOMAIN 136 88	FT	DOMAIN 137 88	FT	DOMAIN 138 88	FT	DOMAIN 139 88	FT	DOMAIN 140 88	FT	DOMAIN 141 88	FT	DOMAIN 142 88	FT	DOMAIN 143 88	FT	DOMAIN 144 88	FT	DOMAIN 145 88	FT	DOMAIN 146 88	FT	DOMAIN 147 88	FT	DOMAIN 148 88	FT	DOMAIN 149 88	FT	DOMAIN 150 88	FT	DOMAIN 151 88	FT	DOMAIN 152 88	FT	DOMAIN 153 88	FT	DOMAIN 154 88	FT	DOMAIN 155 88	FT	DOMAIN 156 88	FT	DOMAIN 157 88	FT	DOMAIN 158 88	FT	DOMAIN 159 88	FT	DOMAIN 160 88	FT	DOMAIN 161 88	FT	DOMAIN 162 88	FT	DOMAIN 163 88	FT	DOMAIN 164 88	FT	DOMAIN 165 88	FT	DOMAIN 166 88	FT	DOMAIN 167 88	FT	DOMAIN 168 88	FT	DOMAIN 169 88	FT	DOMAIN 170 88	FT	DOMAIN 171 88	FT	DOMAIN 172 88	FT	DOMAIN 173 88	FT	DOMAIN 174 88	FT	DOMAIN 175 88	FT	DOMAIN 176 88	FT	DOMAIN 177 88	FT	DOMAIN 178 88	FT	DOMAIN 179 88	FT	DOMAIN 180 88	FT	DOMAIN 181 88	FT	DOMAIN 182 88	FT	DOMAIN 183 88	FT	DOMAIN 184 88	FT	DOMAIN 185 88	FT	DOMAIN 186 88	FT	DOMAIN 187 88	FT	DOMAIN 188 88	FT	DOMAIN 189 88	FT	DOMAIN 190 88	FT	DOMAIN 191 88	FT	DOMAIN 192 88	FT	DOMAIN 193 88	FT	DOMAIN 194 88	FT	DOMAIN 195 88	FT	DOMAIN 196 88	FT	DOMAIN 197 88	FT	DOMAIN 198 88	FT	DOMAIN 199 88	FT	DOMAIN 200 88	FT	DOMAIN 201 88	FT	DOMAIN 202 88	FT	DOMAIN 203 88	FT	DOMAIN 204 88	FT	DOMAIN
----	----------------------	----	--------------------------	----	----------------------------	----	------------------------	----	---------------------------	----	------------------------	----	-----------------------------------	----	-----------------------------------	----	---------------------------	----	-----------------------------------	----	-----------------------------------	----	--	----	-------------	----	-------------	----	-------------	----	-------------	----	-------------	----	-------------	----	-------------	----	-------------	----	-------------	----	--------------	----	--------------	----	--------------	----	--------------	----	--------------	----	--------------	----	--------------	----	--------------	----	--------------	----	--------------	----	--------------	----	--------------	----	--------------	----	--------------	----	--------------	----	--------------	----	--------------	----	--------------	----	--------------	----	--------------	----	--------------	----	--------------	----	--------------	----	--------------	----	--------------	----	--------------	----	--------------	----	--------------	----	--------------	----	--------------	----	--------------	----	--------------	----	--------------	----	--------------	----	--------------	----	--------------	----	--------------	----	--------------	----	--------------	----	--------------	----	--------------	----	--------------	----	--------------	----	--------------	----	--------------	----	--------------	----	--------------	----	--------------	----	--------------	----	--------------	----	--------------	----	--------------	----	--------------	----	--------------	----	--------------	----	--------------	----	--------------	----	--------------	----	--------------	----	--------------	----	--------------	----	--------------	----	--------------	----	--------------	----	--------------	----	--------------	----	--------------	----	--------------	----	--------------	----	--------------	----	--------------	----	--------------	----	--------------	----	--------------	----	--------------	----	--------------	----	--------------	----	--------------	----	--------------	----	--------------	----	--------------	----	--------------	----	--------------	----	--------------	----	--------------	----	--------------	----	--------------	----	--------------	----	--------------	----	--------------	----	---------------	----	---------------	----	---------------	----	---------------	----	---------------	----	---------------	----	---------------	----	---------------	----	---------------	----	---------------	----	---------------	----	---------------	----	---------------	----	---------------	----	---------------	----	---------------	----	---------------	----	---------------	----	---------------	----	---------------	----	---------------	----	---------------	----	---------------	----	---------------	----	---------------	----	---------------	----	---------------	----	---------------	----	---------------	----	---------------	----	---------------	----	---------------	----	---------------	----	---------------	----	---------------	----	---------------	----	---------------	----	---------------	----	---------------	----	---------------	----	---------------	----	---------------	----	---------------	----	---------------	----	---------------	----	---------------	----	---------------	----	---------------	----	---------------	----	---------------	----	---------------	----	---------------	----	---------------	----	---------------	----	---------------	----	---------------	----	---------------	----	---------------	----	---------------	----	---------------	----	---------------	----	---------------	----	---------------	----	---------------	----	---------------	----	---------------	----	---------------	----	---------------	----	---------------	----	---------------	----	---------------	----	---------------	----	---------------	----	---------------	----	---------------	----	---------------	----	---------------	----	---------------	----	---------------	----	---------------	----	---------------	----	---------------	----	---------------	----	---------------	----	---------------	----	---------------	----	---------------	----	---------------	----	---------------	----	---------------	----	---------------	----	---------------	----	---------------	----	---------------	----	---------------	----	---------------	----	---------------	----	---------------	----	---------------	----	---------------	----	---------------	----	---------------	----	---------------	----	---------------	----	---------------	----	--------

358 QOPAGHCRIVREBIFESYRQIMKMRPKDKKGLVYKFRGEBGLDYGGVAREWLYLL 417
 RA Huan Z., Zavelan M., Zhu Y., Zimmer A., Carninci P., Hayatsu N.,
 RA Hirozane-Kishikawa T., Kono H., Nakamura M., Sakazume N., Sato K.,
 449 -RILPGQCHIKVRKNIFEDAYQEIINROTPEJDKGLMIKFDGEEGLDYGGVAREWLYLL 507
 RA Shiraki T., Waki K., Kawai J., Aizawa K., Arakawa T., Fukuda S.,
 RA Hara A., Hashizume W., Imotani K., Ishii Y., Itoh M., Kagawa I.,
 RA Miyazaki A., Sakai K., Sasaki D., Shibata K., Shinagawa A.,
 RA Yasunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J.,
 RA Birney E., Havaehizaki Y.;
 RA "Analysis of the mouse transcriptome based on functional annotation of
 RT 60,770 full-length cDNAs.";
 RT Nature 420:563-573(2002).
 RN [3]
 RN FUNCTION, AND INTERACTION WITH JUN AND JUNB.
 RP MEDLINE=21864584; PubMed=11828324;
 RX Fang D., Elly C., Gao B., Fang N., Altman Y., Joazeiro C., Hunter T.,
 RA Copeland N., Jenkins N., Liu Y.C.;
 RA "Dysregulation of T lymphocyte function in itchy mice: a role for
 RT Itch in Th2 differentiation.";
 RT Nat. Immunol. 3:281-287(2002).
 RN [4]
 RN INTERACTION WITH NOTCH1, AND MUTAGENESIS OF CYS-832.
 RP MEDLINE=20549573; PubMed=10940313;
 RX Qiu L., Joazeiro C., Fang N., Wang H.-Y., Elly C., Altman Y., Fang D.,
 RA Hunter T., Liu Y.-C.;
 RA "Recognition and ubiquitination of Notch by Itch, a hec-type E3
 RT ubiquitin ligase.";
 RT J. Biol. Chem. 275:35734-35737(2000).
 RN [5]
 RN INTERACTION WITH OCNL.
 RP MEDLINE=21895815; PubMed=11782481;
 RX Traweger A., Fang D., Liu Y.-C., Stelzhammer W., Krizbai I.A.,
 RA Fresser F., Bauer H.-C., Bauer H.;
 RA "The tight junction-specific protein occludin is a functional target
 RT of the E3 ubiquitin-protein ligase Itch.";
 RT J. Biol. Chem. 277:10201-10208(2002).
 CC -!- FUNCTION: E3 ubiquitin-protein ligase which accepts ubiquitin from
 CC an E2 ubiquitin-conjugating enzyme in the form of a thioester and
 CC then directly transfers the ubiquitin to targeted substrates.
 CC Regulates the transcriptional activity of several transcription of
 CC factors, and probably plays an important role in the regulation of
 CC immune response.
 CC -!- PATHWAY: Ubiquitin conjugation; third step.
 CC -!- SUBUNIT: Interacts via its WW domains with OCLN, NOTCH1, JUN and
 CC JUNB. Interacts with DRPLA and NFE2 (By similarity).
 CC -!- SUBCELLULAR LOCATION: Nuclear and cytoplasmic.
 CC -!- ALTERNATIVE PRODUCTS:
 CC Event-Alternative splicing; Named isoforms=2;
 CC Name=1;
 CC IsoId=Q8C863-1; Sequence=Displayed;
 CC Note=Major form;
 CC Name=2;
 CC IsoId=Q8C863-2; Sequence=VSP_008452, VSP_008453;
 CC -!- TISSUE SPECIFICITY: Widely expressed.
 CC -!- PTM: Phosphorylated on tyrosine residues (By similarity).
 CC -!- DISEASE: Defects in Itch are the cause of the itchy phenotype
 CC which is an inflammatory and immunological condition characterized
 CC by inflammation in the lung and stomach, hyperplasia in lymphoid
 CC and hematopoietic cells and constant itching in the skin.
 CC -!- SIMILARITY: Contains 1 C2 domain.
 CC -!- SIMILARITY: Contains 1 HECT-type E3 ubiquitin-protein ligase
 CC domain.
 CC -!- SIMILARITY: Contains 4 WW domains.
 CC -!- CAUTION: It is uncertain whether Met-1 or Met-11 is the initiator.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>)
 CC or send an email to license@isb-sib.ch.
 CC EMBL; AF037454; AAB99764.1; ALT_INIT.
 DR

358 QOPAGHCRIVREBIFESYRQIMKMRPKDKKGLVYKFRGEBGLDYGGVAREWLYLL 417
 RA Huan Z., Zavelan M., Zhu Y., Zimmer A., Carninci P., Hayatsu N.,
 RA Hirozane-Kishikawa T., Kono H., Nakamura M., Sakazume N., Sato K.,
 449 -RILPGQCHIKVRKNIFEDAYQEIINROTPEJDKGLMIKFDGEEGLDYGGVAREWLYLL 507
 RA Shiraki T., Waki K., Kawai J., Aizawa K., Arakawa T., Fukuda S.,
 RA Hara A., Hashizume W., Imotani K., Ishii Y., Itoh M., Kagawa I.,
 RA Miyazaki A., Sakai K., Sasaki D., Shibata K., Shinagawa A.,
 RA Yasunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J.,
 RA Birney E., Havaehizaki Y.;
 RA "Analysis of the mouse transcriptome based on functional annotation of
 RT 60,770 full-length cDNAs.";
 RT Nature 420:563-573(2002).
 RN [3]
 RN FUNCTION, AND INTERACTION WITH JUN AND JUNB.
 RP MEDLINE=21864584; PubMed=11828324;
 RX Fang D., Elly C., Gao B., Fang N., Altman Y., Joazeiro C., Hunter T.,
 RA Copeland N., Jenkins N., Liu Y.C.;
 RA "Dysregulation of T lymphocyte function in itchy mice: a role for
 RT Itch in Th2 differentiation.";
 RT Nat. Immunol. 3:281-287(2002).
 RN [4]
 RN INTERACTION WITH NOTCH1, AND MUTAGENESIS OF CYS-832.
 RP MEDLINE=20549573; PubMed=10940313;
 RX Qiu L., Joazeiro C., Fang N., Wang H.-Y., Elly C., Altman Y., Fang D.,
 RA Hunter T., Liu Y.-C.;
 RA "Recognition and ubiquitination of Notch by Itch, a hec-type E3
 RT ubiquitin ligase.";
 RT J. Biol. Chem. 275:35734-35737(2000).
 RN [5]
 RN INTERACTION WITH OCNL.
 RP MEDLINE=21895815; PubMed=11782481;
 RX Traweger A., Fang D., Liu Y.-C., Stelzhammer W., Krizbai I.A.,
 RA Fresser F., Bauer H.-C., Bauer H.;
 RA "The tight junction-specific protein occludin is a functional target
 RT of the E3 ubiquitin-protein ligase Itch.";
 RT J. Biol. Chem. 277:10201-10208(2002).
 CC -!- FUNCTION: E3 ubiquitin-protein ligase which accepts ubiquitin from
 CC an E2 ubiquitin-conjugating enzyme in the form of a thioester and
 CC then directly transfers the ubiquitin to targeted substrates.
 CC Regulates the transcriptional activity of several transcription of
 CC factors, and probably plays an important role in the regulation of
 CC immune response.
 CC -!- PATHWAY: Ubiquitin conjugation; third step.
 CC -!- SUBUNIT: Interacts via its WW domains with OCLN, NOTCH1, JUN and
 CC JUNB. Interacts with DRPLA and NFE2 (By similarity).
 CC -!- SUBCELLULAR LOCATION: Nuclear and cytoplasmic.
 CC -!- ALTERNATIVE PRODUCTS:
 CC Event-Alternative splicing; Named isoforms=2;
 CC Name=1;
 CC IsoId=Q8C863-1; Sequence=Displayed;
 CC Note=Major form;
 CC Name=2;
 CC IsoId=Q8C863-2; Sequence=VSP_008452, VSP_008453;
 CC -!- TISSUE SPECIFICITY: Widely expressed.
 CC -!- PTM: Phosphorylated on tyrosine residues (By similarity).
 CC -!- DISEASE: Defects in Itch are the cause of the itchy phenotype
 CC which is an inflammatory and immunological condition characterized
 CC by inflammation in the lung and stomach, hyperplasia in lymphoid
 CC and hematopoietic cells and constant itching in the skin.
 CC -!- SIMILARITY: Contains 1 C2 domain.
 CC -!- SIMILARITY: Contains 1 HECT-type E3 ubiquitin-protein ligase
 CC domain.
 CC -!- SIMILARITY: Contains 4 WW domains.
 CC -!- CAUTION: It is uncertain whether Met-1 or Met-11 is the initiator.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>)
 CC or send an email to license@isb-sib.ch.
 CC EMBL; AF037454; AAB99764.1; ALT_INIT.
 DR

RA Huckle E., Hunt A.R., Hunt S.E., Jekosch K., Johnson C.M., Johnson D.,
 RA Kay M.P., Kimberley A.M., King A., Knighes A., Laird G.K., Lawlor S.,
 RA Lehtela M.H., Leversha M., Lloyd C., Lloyd D.M., Lovell J.D.,
 RA Marsh V.L., Martin S.L., McConachie L.J., McLay K., McMurray A.A.,
 RA Milne S., Mistry D., Moore M.J.P., Mullikin J.C., Nickerson T.,
 RA Oliver K., Parker A., Patel R., Pearce T.A.V., Peck A.I.,
 RA Phillimore B.J.C.T., Prathalingam S.R., Plumb R.W., Ramsay H.,
 RA Rice C.M., Ross M.T., Scott C.E., Sehra H.K., Showkeen R., Sims S.,
 RA Stuce C.D., Smith M.L., Soderlund C., Steward C.A., Sulston J.E.,
 RA Swann M., Sycamore N., Taylor R., Tee L., Thomas D.W., Thorpe A.,
 RA Tracey A.C., Tromans A.C., Vaudin M., Wall M., Wallis J.M.,
 RA Whitehead S.L., Whittaker P., Willey D.L., Williams L., Williams S.A.,
 RA Wilming L., Wray P.W., Hubbard T., Durbin R.M., Bentley D.R., Beck S.,
 RA Rogers J.,
 RT "The DNA sequence and comparative analysis of human chromosome 20";
 RL Nature 414:865-871 (2001).
 [4]
 RN SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).
 RC TISSUE=Kidney, and placenta;
 RX MEDLINE=22398257; PubMed=12477932;
 RA Klausner R.D., Collins F.S., Wagner L., Shennan C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buettow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavanto T.L., Scheetz T.E.,
 RA Brownstein M.J., Uedin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Iqbal N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.G., Hulyk S.W.,
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahy J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butlerfield Y.S.N., Krzywinski M.I., Skalska U., Smalhus D.E.,
 RA Schnerch A., Schein J.B., Jones S.J.M., Marra M.A.,
 RT "Generation and initial analysis of more than 15,000 full-length human
 RT and mouse cDNA sequences";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
 [5]
 RN SEQUENCE OF 83-903 FROM N.A. (ISOFORM 2), AND INTERACTION WITH DRPLA.
 RP TISSUE=Fetal brain;
 RX MEDLINE=98313405; PubMed=9647693;
 RA Wood J.D., Yuan J., Margolis R.L., Colomer V., Duan K., Kushi J.,
 RA Kaminsky Z., Kleiderlein J.J., Sharp A.H., Ross C.A.;
 RT "Atrophin-1, the DRPLA gene product, interacts with two families of WW
 RT domain-containing proteins.";
 RL Mol. Cell. Neurosci. 11:149-160 (1998).
 [6]
 RN SEQUENCE OF 463-470; 503-510; 514-526; 644-665 AND 875-881,
 RP INTERACTION WITH LMP2A, AND MUTAGENESIS OF CYS-871.
 RX TISSUE=B-cell;
 RA Winberg G., Mateskova L., Chen F., Plant P., Rotin D., Gish G.,
 RA Ingham R., Erbenberg I., Pauson T.,
 RT "Latent membrane protein 2A of Epstein-Barr virus binds WW domain E3
 RT protein-ubiquitin ligases that ubiquitinate B-cell tyrosine kinases.";
 RL Mol. Cell. Biol. 20:8526-8535 (2000).
 [7]
 RN INTERACTION WITH CBL, AND PHOSPHORYLATION.
 RP MEDLINE=22323223; PubMed=12226085;
 RA Courbado J.-R., Fiore F., Adelaide J., Borg J.P., Birnbaum D.,
 RA Oltendorff V.;
 RT "Interaction between two ubiquitin-protein isopeptide ligases of
 RT different classes, CBL and AIP4/ITC.";
 RL J. Biol. Chem. 277:45267-45275 (2002).
 CC -!- FUNCTION: E3 ubiquitin-protein ligase which accepts ubiquitin from
 CC an E2 ubiquitin-conjugating enzyme in the form of a thioester and
 CC then directly transfers the ubiquitin to targeted transcription
 CC regulates the transcriptional activity of several transcription
 CC factors, and probably plays an important role in the regulation of
 CC immune response. Downregulates Epstein-Barr virus LMP2A activity

CC in B cell signaling.
 CC -!- PATHWAY: Ubiquitin conjugation; third step.
 CC -!- SUBUNIT: Interacts via its WW domains with DRPLA, NFE2 and CBL.
 CC Interacts with Epstein-Barr virus LMP2A. Interacts with NOTCH1,
 CC OCIN, JUN and JUNB (By similarity).
 CC -!- SUBCELLULAR LOCATION: Nuclear and cytoplasmic (By similarity).
 CC -!- ALTERNATIVE PRODUCTS:
 CC Event=Alternative splicing; Named isoforms=2;
 CC Name=1;
 CC IsoId=Q96J02-1; Sequence=Displayed;
 CC Note=No experimental confirmation available;
 CC Name=2;
 CC IsoId=Q96J02-2; Sequence=VSP_008451;
 CC TISSUE SPECIFICITY: Widely expressed.
 CC -!- PTM: Phosphorylated on tyrosine residues.
 CC -!- SIMILARITY: Contains 1 C2 domain.
 CC -!- SIMILARITY: Contains 1 HECT-type E3 ubiquitin-protein ligase
 CC domain.
 CC -!- SIMILARITY: Contains 4 WW domains.
 CC -!- CAUTION: Ref.3 sequence differs from that shown due to erroneous
 CC gene model prediction.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
 CC or send an email to license@sib-sib.ch).
 CC -----
 CC EMBL; AF095745; AAK39399.1; -;
 CC EMBL; AB066663; BAB39389.1; -;
 CC EMBL; AL109923; CAC09387.2; -;
 CC EMBL; AL356299; CAC09530.1; ALT_SEQ.
 CC EMBL; BC006848; AAH06848.1; -;
 CC EMBL; BC011571; AAH11571.1; -;
 CC EMBL; AF038564; AAC04845.1; -;
 CC Genew; HGNC:13890; ITC.
 CC MIM; 606409; -;
 CC InterPro; IPR000569; HECT_domain.
 CC Pfam; PF06632; HECT; 1.
 CC PROSITE; PS00499; C2 DOMAIN 1; FALSE_NEG.
 CC PROSITE; PS00004; C2 DOMAIN 2; 1.
 CC PROSITE; PS0237; HECT; 1.
 CC PROSITE; PS01159; WW DOMAIN 1; 4.
 CC PROSITE; PS00020; WW DOMAIN 2; 4.
 CC Ub1 conjugation pathway; Ligase; Nuclear protein; Repeat;
 KW Phosphorylation; Alternative splicing.
 FT DOMAIN 5 99
 FT C2 DOMAIN.
 FT WW 1.
 FT DOMAIN 326 359
 FT WW 2.
 FT DOMAIN 359 391
 FT WW 3.
 FT DOMAIN 438 471
 FT WW 4.
 FT DOMAIN 478 511
 FT HECT.
 FT DOMAIN 589 603
 FT POLY-PRO.
 FT DOMAIN 259 264
 FT UBIQUITIN (BY SIMILARITY).
 FT BINDING 871 871
 FT NGVSLCLPRLCNSAHSALCNLCPLGSLDPSISASRVAGFT
 FT G -> S (in isoform 2).
 FT /FTID=VSP_008451.
 FT C->A: LOSS OF UBIQUITIN PROTEIN LIGASE
 FT MUTAGEN 871 871
 FT ACTIVITY.
 FT SEQUENCE 903 AA; 102802 MW; 6777A2043C7B67BC CRC64;
 SQ
 Query Match 33.2%; Score 1288.5; DB 1; Length 903;
 Best Local Similarity 33.8%; Pred. No. 2.9e-86;
 Matches 312; Conservative 128; Mismatches 239; Indels 245; Gaps 29;
 QY 6 KRLTLVLCNAK-NLAKDSEFLPDPEAKIVDVGSGGCHSTDTVNTLDPKMNQHYDLYGK 64
 DB 19 QLQITVISAKLXENKKNWFG-PSPIYEVTVTD--GOSKTEKCNNTNSPAWQPLTVITVP 75
 QY 65 TDSITISVNHKKIKH-----KQAGFLG-- 88


```
Db 76 VSKLHFRVWSHTLSDVLLGTAALDIYETLKSNNMKLEEVVVVTLQGGDKPTETIGDL 135
QY 89 --CV-----RLSNALRLKDTGY-----RLDL-----CKL-----NPSDITDAV 121
Db 136 SICLDGLQLESEWINGETTCSENGVSLCLPRLCNSAISAHNCNLCPLGSLSDSPISASRV 195
QY 122 RQIVVUS-----LQTFD--RITGGS-----VVDCEGL-----LENGTGYEDSGP 160
Db 196 AGFTGASQNDGSRKDETRVSTNGSDPDAGAGENRVSGNPFSLNNG--FKPBRP 253
QY 161 GRPLSCFMBEAPYTDSTGAAGGNCRFVES-----DQGRVYYVDHVEKRTTWDPR-- 358
Db 254 PRP-----SRPPPTPRPASVNGSPSATSESDSGSTGSLPPTNTNTSEGATSGLLIP 308
QY 193 -----PSQORLOAQLRNPVGRSL-----QTQWRPHGH 223
Db 309 LPISGSGRPLNPVTOAPLPGEWQRV-----DQGRVYYVDHVEKRTTWDPR-- 358
QY 224 QSPPELPEGEQRTTQGGVYFHTQTGSTWHDPRIP-----PSEGSLEDEELPAQRYERD 343
Db 471 -SQQLNE--KLPPEGWEMRFTVDGIPYVDHNRRTTYIDPTGKSALDNGPQIAYVRD 527
QY 344 LVQKLVLR---HELSLOQPQAGHCRIEVSREIFEESYRQIMQVRPKDLKRLMKVFRG 400
Db 528 FXAKVQYFRWCQQLAMPO---HIKITVTRKTLFEDSFQOIMSPSQDRLRLWVIFPG 583
QY 401 ERLDYGGVAREWLVLCHEMLNPVYGLFOYSTDNIYMLQINPDSINPDHLSYFHFVGR 460
Db 584 ERLDYGGVAREWLVLCHEMLNPVYGLFOYSTDNIYMLQINPDSINPDHLSYFHFVGR 460
QY 461 INGLAVFHGYINGFTVPFYKQLCKPTQLSDLESVDPELHKSLSVLENDITPV-LDH 519
Db 644 FIATVALFHGKFTDGTSPFYKRLNKPVLGKDLSDIDPEFFNSLIWKENNIEBCDLEM 703
QY 520 TFCVHNAPGRILQHLKPNRNPVTEENKEYVRLYVNRPMFGIEAQFALQKNGNE 579
Db 704 YFSVDKEILGETKSHDLKPNCGNLIYVTEENKEYIRMVAVELSLRSGVEEQTQAFEGFNE 763
QY 580 LIPQHLKPFDOKELELIGGLDKIDLNDKSNTRLKHCVADSNIVRMFWQAVETFDER 639
Db 764 ILPQVLYQVDAKELEVLICGQEDLDNDQORHAIYRHVARTSKQIMFWQFVKEIDNEK 823
QY 640 RARLLQFVTGSTRVPLQGFALQGSTGAAPRLFTIHLIDANTDNLKPAHTCFNRIDIPP 699
Db 824 RMELLQFVTGTRKLPVGGFADLMGNS--GPKQFCIEKV-GKENWLPRSHTCFNRDLDP 879
QY 700 YESYKLYBKLTAVEETCGFAVE 723
Db 880 YKSYSQLEKLLFAETEETGFGQE 903

RESULT 9
ID NED4 HUMAN STANDARD; PRT; 1000 AA.
AC P46334;
DT 01-NOV-1995 (Rel. 32, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Ubiquitin-protein ligase Nedd-4 (EC 6.3.2.-).
GN NEDD4 OR KIAA0093.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
```

```
RP RA Hinz U.;
RN Unpublished observations (JUN-2003).
RC [2]
RX SEQUENCE OF 74-1000 FROM N.A.
RC TISSUE=Bone marrow;
RX MEDLINE=95308325; PubMed=7785527;
RA Nagase T., Miyajima N., Tanaka A., Suzuki T., Seki N., Sato S.,
RA Tabata S., Ishikawa K.-I., Kawabayashi Y., Kotani H., Nomura N.;
RT "Prediction of the coding sequences of unidentified human genes. III.
RT The coding sequences of 40 new genes (KIAA0081-KIAA0120) deduced by
RT analysis of cDNA clones from human cell line KG-1.";
RL DNA Res. 2:37-43(1995).
CC -!- FUNCTION: E3 ubiquitin-protein ligase which accepts ubiquitin from
CC an E2 ubiquitin-conjugating enzyme in the form of a thioester and
CC then directly transfers the ubiquitin to targeted substrates (By
CC similarity).
CC -!- PATHWAY: Ubiquitin conjugation; Third step.
CC -!- SUBUNIT: Interacts with UBE2D2. Binds SCNN1A, SCNN1B and SCNN1G
CC (By similarity).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -!- MISCELLANEOUS: A cysteine residue is required for ubiquitin-
CC thiolester formation.
CC -!- SIMILARITY: Contains 1 C2 domain.
CC -!- SIMILARITY: Contains 4 WW domains.
CC -!- SIMILARITY: Contains 1 HECT-type E3 ubiquitin-protein ligase
CC domain.
CC -!- CAUTION: The sequence of the N-terminus was deduced from the
CC genomic sequence.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AC039057; -; NOT ANNOTATED_CDS.
CC EMBL; D42055; BAA07655.1; -;
CC HSSP; Q13526; 1PIN.
CC Genew; HGNC:7727; NEDD4.
CC MIN; 602278; -.
CC InterPro; IPR000008; C2.
CC InterPro; IPR008973; C2 CaLB.
CC InterPro; IPR000569; HECT domain.
CC InterPro; IPR001202; WW_Rap5_WMP.
CC Pfam; PF00168; C2; 1.
CC Pfam; PF00632; HECT; 1.
CC Pfam; PF00397; WW; 4.
CC SMART; SM00239; C2; 1.
CC SMART; SM00119; HECTC; 1.
CC SMART; SM00456; WW; 4.
CC PROSITE; PS00499; C2_DOMAIN_1; 1.
CC PROSITE; PS00004; C2_DOMAIN_2; 1.
CC PROSITE; PS00237; HECT; 1.
CC PROSITE; PS01159; WW_DOMAIN_1; 4.
CC PROSITE; PS00020; WW_DOMAIN_2; 4.
CC Ubl conjugation pathway; Ligase; Repeat.
KW DOMAIN 9 44
FT DOMAIN 77 81 POLY-SER.
FT DOMAIN 106 207 C2 DOMAIN.
FT DOMAIN 285 288 POLY-GLN.
FT DOMAIN 291 324 WW 1.
FT DOMAIN 448 481 WW 2.
FT DOMAIN 521 554 WW 3.
FT DOMAIN 573 606 WW 4.
FT DOMAIN 565 1000 HECT.
FT BINDING 967 967 UBIQUITIN (BY SIMILARITY).
SQ SEQUENCE 1000 AA; 114936 MW; 372808850C149CB CRC64;
Query Match 32.4%; Score 1259; DB 1; Length 1000;
Best Local Similarity 33.6%; Pred. No. 4.9e-84;
```


RX MEDLINE=22338357; PubMed=12450395;
RA Galinier R., Gout E., Iortat-Jacob H., Wood J., Chroboczek J.;
RT "Adenovirus protein involved in virus internalization recruits
RL Biochemistry 41:14299-14305(2002).
[7]
RP X-RAY CRYSTALLOGRAPHY (2.2 ANGSTROMS) OF 546-917, MUTAGENESIS OF
RP GLU-614; HIS-621; ASP-675; GLU-799; MET-804; ARG-845; GLN-848
RP AND ARG-855, AND FUNCTION.
RX MEDLINE=22423789; PubMed=12535537;
RA Verdecia M.A., Jazeiro C.A.P., Wells N.J., Ferrer J.-L., Bowman M.E.,
RA Hunter T., Noel J.P.;
RT "Conformational flexibility underlies ubiquitin ligation mediated by
RT the WPI1 HECT domain E3 ligase.";
RL Mol. Cell 11:249-259(2003).
CC -!- FUNCTION: E3 ubiquitin-protein ligase which accepts ubiquitin from
CC an E2 ubiquitin-conjugating enzyme in the form of a thioester and
CC then directly transfers the ubiquitin to targeted substrates.
CC -!- PATHWAY: Ubiquitin conjugation; third step.
CC -!- SUBUNIT: Binds KLF2 (By similarity). Binds SCNN1A, SCNN1B, SCNN1G,
CC WBP1, WBP2, DRPLA and adenovirus type 2 PIII.
CC -!- ALTERNATIVE PRODUCTS:
CC Event-Alternative splicing: Named isoforms=6;
CC Comment=Additional isoforms seem to exist;
CC Name=1; Synonyms=A;
CC IsoId=Q9H0M0-1; Sequence=VSP_007601, VSP_007603;
CC Name=2; Synonyms=B;
CC IsoId=Q9H0M0-2; Sequence=VSP_007601, VSP_007602;
CC Name=3; Synonyms=C;
CC IsoId=Q9H0M0-3; Sequence=VSP_007602;
CC Name=4; Synonyms=D;
CC IsoId=Q9H0M0-4; Sequence=Not described;
CC Name=5; Synonyms=E;
CC IsoId=Q9H0M0-5; Sequence=Not described;
CC Name=6; Synonyms=F;
CC IsoId=Q9H0M0-6; Sequence=VSP_007600;
CC -!- TISSUE SPECIFICITY: Detected in heart, placenta, pancreas, kidney,
CC liver, skeletal muscle, bone marrow, fetal brain, and at much
CC lower levels in adult brain and lung. Isoforms 1 and 5 predominate
CC in all tissues tested, except in testis and bone marrow, where
CC isoform 5 is expressed at much higher levels than isoform 1.
CC -!- MISCELLANEOUS: A cysteine residue is required for ubiquitin-
CC thiolester formation.
CC -!- SIMILARITY: Contains 1 C2 domain.
CC -!- SIMILARITY: Contains 4 WW domains.
CC -!- SIMILARITY: Contains 1 HECT-type E3 ubiquitin-protein ligase
CC domain.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

DR InterPro: IPR000008; C2.
DR InterPro: IPR008973; C2_CalB.
DR InterPro: IPR000569; HECT_domain.
DR InterPro: IPR02349; WW.
DR InterPro: IPR011202; WW_Rsp5_WWP.
DR Pfam: PF00168; C2; 1.
DR Pfam: PF00632; HECT; 1.
DR Pfam: PF00397; WW; 4.
DR PRINTS: PR00403; WWDOMAIN.
DR SMART: SM00239; C2; 1.
DR SMART: SM00119; HECTC; 1.
DR SMART: SM00456; WW; 4.
DR PROSITE: PS00499; C2_DOMAIN_1; FALSE_NEG.
DR PROSITE: PS00004; C2_DOMAIN_2; 1.
DR PROSITE: PS00237; HECT; 1.
DR PROSITE: PS01159; WW_DOMAIN_1; 4.
DR PROSITE: PS00020; WW_DOMAIN_2; 2.
KW Ub1 conjugation pathway; Ligase; Repeat; Alternative splicing;
KW 3D-structure.
FT DOMAIN 5 98 C2 DOMAIN.
FT DOMAIN 349 382 WW 1.
FT DOMAIN 381 414 WW 2.
FT DOMAIN 456 489 WW 3.
FT DOMAIN 496 529 WW 4.
FT DOMAIN 588 922 HECT.
FT BINDING 890 890 UBIQUITIN (BY SIMILARITY).
FT VARSPPLIC 23 240 Missing (in isoform 6).
FT VARSPPLIC 112 120 LERVKQLK -> CWLKARME (in isoform 2).
FT VARSPPLIC 112 242 LERVKQLKSLNKGKIAOTGLTVVLDGLVIEQENITNC
FT LERVKQLKSLNKGKIAOTGLTVVLDGLVIEQENITNC
FT SSPTIEIQNGDHALHENGEPSTARTTARLAVEGTNGIDNHV
FT PTSLTVQNSCCSYVNGDNTFSSPSQVAARPKPTPAKPLA
FT SEPADDTV -> F (in isoform 3).
FT VARSPPLIC 121 922 Missing (in isoform 2).
FT VARSPPLIC 614 614 E-A: REDUCES UBIQUITIN TRANSFER.
FT MUTAGEN 621 621 H-A: STRONGLY REDUCES UBIQUITIN
FT MUTAGEN 675 675 D-A: REDUCES UBIQUITIN TRANSFER.
FT MUTAGEN 798 798 E-A: REDUCES UBIQUITIN TRANSFER;
FT MUTAGEN 804 804 M-P: STRONGLY REDUCES UBIQUITIN
FT MUTAGEN 806 806 TRANSFER; WHEN ASSOCIATED WITH P-806.
FT MUTAGEN 845 845 E-S: STRONGLY REDUCES UBIQUITIN
FT MUTAGEN 848 848 TRANSFER; WHEN ASSOCIATED WITH P-804.
FT MUTAGEN 848 848 R-A: NO EFFECT.
FT MUTAGEN 848 848 Q-A: ABOLISHES UBIQUITIN TRANSFER; WHEN
Query Match 31.9%; Score 1239; DB 1; Length 922;
Best Local Similarity 32.5%; Pred.No.1.3e-82;
Matches 300; Conservative 138; Mismatches 282; Indels 224; Gaps 25;
QY 5 IKRLTVLCARNAKQDFRLPDPFAKIVVDGSGQCHSTDTVK--NTLDPKNNQHYDLYV 62
DB 18 LQLQVTVSSAKLRKKNWFGTA-IYTEVVDG--BITKTAKSSSSNPKWDELTVNV 72
QY 63 KGTDSIITISWNHKKIHKQAGELG-----CVLLSNALSRLKD----- 102
DB 73 TPQTTLEFQWNSHRTL---KADALLGKATIDLKQALLIHRKLERVKEQKLSLENKGI 129
QY 103 --TG-----YORLDLCKINPS-----DTDVAVGQIVVSLQTRDRI---GTGG- 139
DB 130 AQTGLTIVLDGLVIEQENITNCSSPTIEIQNGDHALHENGEPSTARTTARLAVEGTNGI 189
QY 140 -SVVDCRGLLENGEVTVDGSGRPLSCFM-----EEPAP-----Y 174
DB 190 DNHVPTSTLVQNSCCSYVNGDNTFSSPSQVAARPKNTPAKPLASEPADDTVNGESSSF 249
QY 175 TDSITCAAGG-----GNC-----RFVESPSQDQORQAQR 203

RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan K., Moore T., Max S.I., Wang J., Heide F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raba S.S., Lequellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyx S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Kettner M., Madan A., Rodriguez S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butterfield V.S.N., Krzyzanski M.T., Skalska U., Smalls D.B.,
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.,
 RA RT "Generation and initial analysis of more than 15,000 full-length
 RL human and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RP [3]
 RP INTERACTION WITH KLF2.
 RX MEDLINE=2136920; PubMed=11375995;
 RA Konkrigt M.D., Wani M.A., Lingrei J.B.;
 RT "Lung Kruessel-like factor contains an autoinhibitory domain that
 RT regulates its transcriptional activation by binding Wp1, an E3
 RT ubiquitin ligase.";
 RL J. Biol. Chem. 276:29299-29306(2001).
 CC -!- FUNCTION: E3 ubiquitin-protein ligase which accepts ubiquitin from
 CC an E2 ubiquitin-conjugating enzyme in the form of a thioester and
 CC then directly transfers the ubiquitin to targeted substrates (by
 CC similarity).
 CC -!- PATHWAY: Ubiquitin conjugation; third step.
 CC -!- SUBUNIT: Binds SCNN1A, SCNN1B, SCNN1G, WBP1, WBP2, DRPLA and
 CC adenovirus type 2 pIII (By similarity). Binds KLF2.
 CC -!- MISCELLANEOUS: A cysteine residue is required for ubiquitin-
 CC thiolester formation (by similarity).
 CC -!- SIMILARITY: Contains 1 C2 domain.
 CC -!- SIMILARITY: Contains 4 WW domains.
 CC -!- SIMILARITY: Contains 1 HECT-type E3 ubiquitin-protein ligase
 CC domain.
 CC -!- CAUTION: Ref.1 (BAC38473) sequence differs from that shown after
 CC position 391; the second part of that clone maps to another
 CC chromosome.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; AK031138; BAC28168.1; -;
 CC EMBL; AK082346; BAC38473.1; ALT SEQ.
 CC EMBL; BC021470; AAH21470.1; ALT_INIT.
 CC EMBL; BC051405; AAH51405.1; -;
 CC MGD; MGI:1861728; Wp1.
 CC GO; GO:0000151; C:ubiquitin ligase complex; ISS.
 CC GO; GO:0005515; F:protein binding; IPI.
 CC GO; GO:0004842; F:ubiquitin-protein ligase activity; ISS.
 CC GO; GO:0007417; P:central nervous system development; ISS.
 CC GO; GO:0030324; P:lung development; TAS.
 CC GO; GO:0016481; P:negative regulation of transcription; IDA.
 CC GO; GO:0016567; P:protein ubiquitination; ISS.
 CC GO; GO:0007165; P:signal transduction; ISS.
 CC GO; GO:0030217; P:T-cell differentiation; TAS.
 CC GO; GO:0046718; P:viral entry; ISS.
 CC InterPro; IPR000008; C2.
 CC InterPro; IPR008973; C2 CalB.
 CC InterPro; IPR000569; HECT_domain.
 CC InterPro; IPR002349; WW.
 CC InterPro; IPR001202; WW_Rep5_WWP.
 CC Pfam; PF00632; HECT.1.

DR Pfam: PF00397; WW; 4.
 DR PRINTS; PR00403; WWDOMAIN.
 DR SMART; SM00119; HECTc; 1.
 DR SMART; SM00456; WW; 4.
 DR PROSITE; PS00499; C2 DOMAIN 1; FALSE_NEG.
 DR PROSITE; PS00004; C2 DOMAIN_2; 1.
 DR PROSITE; PS02337; HECT; 1.
 DR PROSITE; PS01159; WW_DOMAIN_1; 4.
 DR PROSITE; PS00020; WW_DOMAIN_2; 4.
 KW Ub1 conjugation pathway; Ligase; Repeat.
 FT DOMAIN 5 98
 FT C2 DOMAIN
 FT DOMAIN 345 378 WW 1.
 FT DOMAIN 377 410 WW 2.
 FT DOMAIN 452 485 WW 3.
 FT DOMAIN 452 525 WW 4.
 FT DOMAIN 584 918 HECT.
 FT BINDING 886 886 UBIQUITIN (BY SIMILARITY).
 FT CONFLICT 302 302 V -> L (IN REF. 1; BAC28168).
 SQ SEQUENCE 918 AA; 104693 MW; 01478A3C1CFDAA9 CRC64;
 Query Match 31.7%; Score 1232; DB 1; Length 918;
 Best Local Similarity 32.3%; Pred. No. 4.1e-82;
 Matches 302; Conservative 137; Mismatches 247; Indels 248; Gaps 25;
 QY 5 IKIRLVLCARNLAKKDFRLPDPFAKIVDPGSGQCHSTDTVKNLDPKMNQHYDLVYVK 64
 DB 18 LQKVTSSAKLKKKKWFGTA-IYTEIVD--GEVKTAKSSSSSNPKMDEQLIVNVTP 74
 QY 65 TDSITISWNHKKHKQAGFLGCVRLLSNAISRLKD---TGVRDLCLKLNPSDSDAV 121
 DB 75 QTLEFRVWSHHTL--KADA-----DLGKATVDLKVLLTHNRKL-----EKV 115
 QY 122 RQIVVSLQTRDRI--GTGSGVVDRCGLL-----EN----- 150
 DB 116 KQALKSLKNGKNGIVQTGELIVLDGLVIEQEPVNRSSSPPIIQNGDALHENGDPAT 175
 QY 151 -----EGTVED-----SGP-----GRPLSCFMEEP--- 171
 DB 176 RTTPRLPVEGTIGIDNHVSTNTVVPNSCCSHVANGENTPSSPSQVAARPKAPKPVTS 235
 QY 172 APYDTS----- 177
 DB 236 APTSTVNGESSVLADNTSTMGILLPSEDITSTSNCTSTTQBPVQEPASSEHSECI 295
 QY 178 -TGAAGGCGNCFVSPSQDRLQA--QRLNPD-----VSGSLQT-----PQN 218
 DB 296 PSASAEVGPARSLLTDPDSRNNNSVFDKVRQEGCVPLRPSQSGTNTTEALPSGVBQRK 355
 QY 219 RPHGH-----QSPD-LPEGVGEORTTVQGVYFLHTOTGVSTWHDPRIPDLN 264
 DB 356 DPHGTYTVVDNTRTTTWERFQPLPPGERRVDRGRVYVVDNTRTTTWRPTNESVRN 415
 QY 265 -----SYNCDELGLPFGGWEVRSTVSGRIYFVDHNN 295
 DB 416 FEWQSQORNQLQAGAQFNQRYLYSASMLAAENDPYGPLPFGWEKRVDSITRVYFNHNT 475
 QY 296 RTTQTDRLHIMNHQ-----COLKPSQPLPSPGSL 330
 DB 476 KTTQWEDPTQGLPNEEPLPEGWIRYTRGVRFVDNTRTTTTFKDFRNGKSVTKGG- 534
 QY 331 EDEELPAQRYERDLVQKLKVLRLHLSLQQPOAGHCRIEVSREEIFEESYQIOMKRPDL 390
 DB 535 -----PQIAYERSFRWKLAFERY-LCCSNALPSHVKNVSRQTLFEDSFQIMALKPYDL 588
 QY 391 KRLRMVFRGREGLDYGVVAREWLYLICHEMLNPPYGLFQYSTDNITVQLNPDSINPD 450
 DB 589 RRLYVIFRGEGLDYGLAREWFFLLSHSVINPMYCLFEYAGKKNYCLQINPASTINPD 648
 QY 451 HLSVPHFVGRIMGLAVPHGHVINGFTVPFYKQLLKGKPIQLSDLESVDPELHSLVWILE 510
 DB 649 HLSVFCFGRFIALMFLHGFIDTGFSLPFFYKMLSKKLTKDLESIDTFEYNSLIWRD 708
 QY 511 NDTITV-LDHTTCVEHNAFGRILQHELKPNGRNVPTVEENKSVRLYVNVNWRFMRGIEAQ 569


```

Db 709 NNEECGLEWYSVDMEIIGKVTSHDKLGGNSILVTEENKDEYIGLMTWFSRGVQEQ 768
QY 570 FLALQXGNELIPQHLKPFQKLELEIIGGLDKIDLNDWKSNTLKHCVADSNIVRWFW 629
Db 769 TRAFLDGFEVVPQLQVDFEKELEVMICGQEVDLADQRTVYRHYTRNSKQIWNF 828
QY 630 QAVETFDERRARLLQFVGTGRVPLQGLKALQSGTGAAGPRLFTIHLIDANTDNLPAH 689
Db 829 QFVKETDNRVRLQFVGTGRVPLQGLKALQSGTGAAGPRLFTIHLIDANTDNLPAH 884
QY 690 TCFNRIDIPPSYSEKLYEKLTAVEETCGFAVE 723
Db 885 TCFNRDLDPYKSYEQLEKLLFAIETGFGQE 918

RESULT 13
NED4 RAT
ID NED4 RAT STANDARD; PRT; 887 AA.
AC O62940;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Ubiquitin-protein ligase Nedd-4 (EC 6.3.2.-).
GN NEDB4.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A., AND INTERACTION WITH SCNN1A; SCNN1B AND SCNN1G.
RC TISSUE=Lung;
RX MEDLINE=96221297; PubMed=8665844;
RA Staub O., Dho S., Henry P., Correa J., Ishikawa T., McGlade J.,
RA Rotin D.;
RT "WW domains of Nedd4 bind to the proline-rich PY motifs in the
RT epithelial Na+ channel deleted in Liddle's syndrome.";
RL EMBO J. 15:2371-2380(1996).
RN [2]
RP STRUCTURE BY NMR OF 452-499 IN COMPLEX WITH SCNN1B, AND INTERACTION
RP WITH SCNN1A; SCNN1B AND SCNN1G.
RX MEDLINE=2123577; PubMed=11323714;
RA Kanellis V., Rotin D., Forman-Kay J.D.;
RT "Solution structure of a Nedd4 WW domain-ENAC peptide complex.";
RL Nat. Struct. Biol. 8:407-412(2001).
CC -!- FUNCTION: E3 ubiquitin-protein ligase which accepts ubiquitin from
CC an E2 ubiquitin-conjugating enzyme in the form of a thioester and
CC then directly transfers the ubiquitin to targeted substrates (By
CC similarity).
CC -!- PATHWAY: Ubiquitin conjugation; third step.
CC -!- SUBUNIT: Interacts with UBE2D2 (By similarity). Binds SCNN1A,
CC SCNN1B and SCNN1G.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC -!- TISSUE SPECIFICITY: Ubiquitous. Expression is highest in lung,
CC kidney and brain.
CC -!- MISCELLANEOUS: A cysteine residue is required for ubiquitin-
CC thioester formation.
CC -!- SIMILARITY: Contains 1 C2 domain.
CC -!- SIMILARITY: Contains 3 WW domains.
CC -!- SIMILARITY: Contains 1 HECT-type E3 ubiquitin-protein ligase
CC domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U50842; AAB48949.1; -
CC PIR; S70642; S70642.
CC PDB; 115H; 02-MAY-01.

```

```

DR InterPro; IPR000008; C2.
DR InterPro; IPR00973; C2 CalB.
DR InterPro; IPR000569; HECT domain.
DR InterPro; IPR001202; WW_Rsp5_WWP.
DR Pfam; PF00168; C2; 1.
DR Pfam; PF00632; HECT; 1.
DR Pfam; PF00397; WW; 3.
DR PRINTS; PR00360; C2DOMAIN.
DR SMART; SM00239; C2; 1.
DR SMART; SM00119; HECTC; 1.
DR SMART; SM00456; WW; 3.
DR PROSITE; PS00499; C2 DOMAIN_1; 1.
DR PROSITE; PS50004; C2 DOMAIN_2; 1.
DR PROSITE; PS50237; HECT; 1.
DR PROSITE; PS01159; WW DOMAIN_1; 3.
DR PROSITE; PS50020; WW DOMAIN_2; 3.
KW Ubl conjugation pathway; Ligase; Repeat; 3D-structure.
FT DOMAIN 62 163 C2 DOMAIN.
FT DOMAIN 246 279 WW 1.
FT DOMAIN 402 435 WW 2.
FT DOMAIN 459 492 WW 3.
FT DOMAIN 551 587 HECT.
FT BINDING 854 854 UBIQUITIN (BY SIMILARITY).
SQ SEQUENCE 887 AA; 102394 MW; D74B1097688CD9A1 CRC64;

Query Match 31.6%; Score 1227.5; DB 1; Length 887;
Best Local Similarity 33.9%; Pred.No. 8.3e-82;
Matches 296; Conservative 127; Mismatches 224; Indels 227; Gaps 28;

QY 7 IRLTVLCAKSLAKDFRFLPDPFAKIV-----VPGSGCHSTDTVNTLDPKWN----- 55
Db 77 VRKVIAGIGLAKKILGASDPYRVVTLDPMSGLTSTVQTKTKKSLNPKWNEILFRV 136
QY 56 ---QH-----YDLYVGKTD--SITISVWN-----HKIHKKQG 83
Db 137 LPQHRILFEVDENRLTRDQGLQVDVPLPTENPRMERYTFKDFVLHPRSHKSRV 196
QY 84 AGFLCGVLLLSNAISRLLKDT-----RLKMTYLPKNGSDDENADQAELEFGWVLDQDAATHLQHPPEP 245
Db 197 KGYL-----GY-QLDLG-----KLNPSD--TDVARGQIVV-----SLQTRDRI 135
QY 104 -----GY-QLDLG-----KLNPSD--TDVARGQIVV-----SLQTRDRI 135
Db 246 SPLPFGWEERQDVLGRITYYNHESRTTQWKRPSPEDDLTDENGDIQLQAHGAFTRRQI 305
QY 136 GTGSGVVDRCG-----LLENEGTVYEDSGGRLPS-----C 166
Db 306 SE-----DVGDPDNHESPENWEIVRENTIYSGQAVQSPSPGHPVQVFLABELDRLIT 360
QY 167 FMEEPA---PYTDSTGAAGCGN--CRFVSPSQDQRLQORLNPDPVRSLOTPQNRPH 221
Db 361 MYGNPATSQPVTSNHSRSGSSQTCIFERQPTLPVLLPT-----RIPRDLNS---VNCDE 270
QY 222 GHQSPPELPEGYEQRTTVQGVYFLHTQTGVSTVHDP-----RIPRDLNS---VNCDE 270
Db 401 ---SSGLPPGWEERQDQDRGRSYVDHNSKTTTWSKPTMQDDPRSKIPAHURGTVPDSND 457
QY 271 LGPLPPGWEYRSTVSGRIYFVDHNNRTQFTDRLHHIMNHQCLKEPSQFLPLPSEGL 330
Db 458 LGPLPPGWEERTHTDGRVFFINHNKKTQWEDFRMQNV-----AITGPAEP----- 503
QY 331 EDELPAQRVERDLVKLVLRHLSLQOQQAQCHIEVSREEIFESYQIMKVRPKD- 389
Db 504 -----YSRDYKRYEYFFRKLKKQTDPNKKFKMKURRANILEDYSYRIMGVRADF 554
QY 390 LKRLMKVFRGEGLDYGGVAREWLYLLCHEMLNPPYGLFOYS-TDNIMYLMQINPDSS1- 447
Db 555 LKARLWTEFDGKGLDYGGVAREWFFLISKEMENPYGLFEYATENDNYTLQINPSGLC 614
QY 448 NPDHLSYFHVGRIMGLAVFHGYINGGFTVPYKQLLGPILQISDLESVDPELHSLVW 507
Db 615 NEDHLSYFKFGRVAGNAVYHGKLLDGFIRFPYKMWLQKLITLHDWESVDSEYSSLRW 674

```



```
QY 508 ILENDITFVLDTFCVHNHAFGRILQHELKPNGRNVPTENKKEYVRLVYVNRFRMGIE 567
DB 675 ILENDPTE-LDLRFIIDEELFGQTHLKTGSEVVVTNKNKEVILYVIOWRFYNRQ 733
QY 568 AQFLALOKGNEKLIPQHLKPFDDOKLELLIGGLDKIDLNDKSNTRKHCVA-DSNIVR 626
DB 734 KOMAAFEKGFPELLIPQDLIKFIDNEBELMLCGLGVDVNDWREHTKYKNGYSLNKHQVH 793
QY 627 WFOQAVETFDERRARLLQFVTGSTRVPLQFGKALQGSTGAAGPRFTTHLIDANTDNL 686
DB 794 WFKAVLMDSEKIRLLQFVTGSTRVPMNGFABLYGSLN--GQPSFTVQW-GTDPKLP 849
QY 687 KAHTCFNRIDIPPEVSEKLYELLTAVERTCGF 720
DB 850 RAHTCFNRLLDPPEVSEFDELWDLQMAIENTQGF 883

RESULT 14
WBP2_HUMAN STANDARD; PRT; 870 AA.
AC 000308; Q96C22; Q9BWN6;
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Nedd-4-like ubiquitin-protein ligase WBP2 (EC 6.3.2.-) (WW domain-
DE containing protein 2) (Atropin-1 interacting protein 2) (Aip2).
CN WBP2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A., AND INTERACTION WITH WBP1; WBP2; SCNN1A; SCNN1B
RP AND SCNN1G.
RC TISSUE=Brain, and Placenta;
RX MEDLINE=22388257; PubMed=12477932;
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan B., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettner M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Truchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Groumwood J., Schmutz J., Myers R.M.,
RA Butterfield V.S.N., Krywinski M.I., Skalska U., Smalls D.B.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.,
RA "Generation and initial analysis of more than 15,000 full-length
RA human and mouse cDNA sequences."
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [3]
RP INTERACTION WITH DRPLA, AND TISSUE SPECIFICITY.
RX MEDLINE=98313405; PubMed=9647693;
RA Wood J.D., Yuan J., Margolis R.L., Coloner V., Duan K., Kushi J.,
RA Kaminsky Z., Kleiderlein J.J. Jr., Sharp A.H., Ross C.A.;
RT "Atrophin-1, the DRPLA gene product, interacts with two families of
RT WW domain-containing proteins."
RL Mol. Cell. Neurosci. 11:149-160(1998).
```

```
RN [4]
RP INTERACTION WITH SCNN1A; SCNN1B AND SCNN1G.
RX MEDLINE=22151711; PubMed=12167593;
RA McDonald F.J., Western A.H., McNeill J.D., Thomas B.C., Olson D.R.,
RA Snyder P.M.;
RT "Ubiquitin-protein ligase WBP2 binds to and downregulates the
RT epithelial Na(+) channel."
RL Am. J. Physiol. 283:F431-F436(2002).
RN [5]
RP INTERACTION WITH ADENOVIRUS TYPE 2 PIII.
RX MEDLINE=22338357; PubMed=12450395;
RA Galinier R., Gout E., Lortat-Jacob H., Wood J., Chroboczek J.;
RT "Adenovirus protein involved in virus internalization recruits
RT ubiquitin-protein ligases."
RL Biochemistry 41:14299-14305(2002).
CC -!- FUNCTION: E3 ubiquitin-protein ligase which accepts
CC ubiquitin from an E2 ubiquitin-conjugating enzyme in the form of a
CC thioester and then directly transfers the ubiquitin to targeted
CC substrates (By similarity).
CC -!- PATHWAY: Ubiquitin conjugation; third step.
CC -!- SUBUNIT: Binds SCNN1A, SCNN1B, SCNN1G, WBP1, WBP2, DRPLA and
CC adenovirus type 2 PIII.
CC -!- TISSUE SPECIFICITY: Detected in heart, throughout the brain,
CC placenta, lung, liver, muscle, kidney and pancreas.
CC -!- MISCELLANEOUS: A cysteine residue is required for ubiquitin-
CC thiolester formation.
CC -!- SIMILARITY: Contains 1 C2 domain.
CC -!- SIMILARITY: Contains 4 WW domains.
CC -!- SIMILARITY: Contains 1 HECT-type E3 ubiquitin-protein ligase
CC domain.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; U96114; AAC51325.1; -.
CC DR EMBL; BC000108; AAH00108.1; -.
CC DR EMBL; BC013645; AAH13645.1; -.
CC DR HSSP; Q13526; IPIN.
CC DR MIM; 602308; -.
CC DR GO; GO:0000151; C:ubiquitin ligase complex; TAS.
CC DR GO; GO:0004842; F:ubiquitin-protein ligase activity; TAS.
CC DR GO; GO:0006464; P:protein modification; TAS.
CC DR InterPro; IPR000008; C2.
CC DR InterPro; IPR008973; C2_CaLB.
CC DR InterPro; IPR000569; HECT_domain.
CC DR InterPro; IPR002349; WW.
CC DR Pfam; PF00168; C2; 1.
CC DR Pfam; PF00632; HECT; 1.
CC DR Pfam; PF00397; WW; 4.
CC DR PRINTS; PR00403; WWDOMAIN.
CC DR SMART; SM00233; C2; 1.
CC DR SMART; SM00119; HECTC; 1.
CC DR SMART; SM00456; WW; 4.
CC DR PROSITE; PS00499; C2_DOMAIN_1; FALSE_NEG.
CC DR PROSITE; PS50004; C2_DOMAIN_2; FALSE_NEG.
CC DR PROSITE; PS50237; HECT; 1.
CC DR PROSITE; PS01159; WW_DOMAIN_1; 4.
CC DR PROSITE; PS50020; WW_DOMAIN_2; 4.
KW Ub1 conjugation pathway; Ligase; Repeat.
FT DOMAIN 20 100 C2 DOMAIN.
FT DOMAIN 300 333 WW 1.
FT DOMAIN 330 363 WW 2.
FT DOMAIN 405 437 WW 3.
FT DOMAIN 444 477 WW 4.
FT DOMAIN 536 870 HECT.
FT BLINDING 838 858 UBIQUITIN (BY SIMILARITY).
FT CONFLICT 136 136 E -> K (IN REF. 1).
```

FT	CONFLICT	394	395	SS -> FW (IN REF. 1).	
5Q	SEQUENCE	870 AA;	98911 MW;	FCCD75CBA61F2204 CRC64;	
	Query Match	31.5%;	Score 1225;	DB 1; Length 870;	
	Best Local Similarity	35.5%;	Pred. No. 1.2e-81;		
	Matches	281;	Conservative 115;	Mismatches 227; Indels 169; Gaps 19;	
QY		28	PFKIVVDGSGQCHSTDTVKNTLDPKWNQHY----	DLVGTKDTSITISVNNHKKIHKQG 83	
DB		152	PNGSALTGDS--QLPSRDSGSGTAVAFE--NRHQPPSTNCFGGRS-----	RTHRHSG 198	
QY		84	AGFLGCVRLLSNAISRLKDTGYQRLDLCKLNPSDTPDAVRGQIVVSLQTRD--	IGTGSV 141	
DB		199	AS-----	AKTTPATGEQSPGASRRHQPVKNSGHS 228	
QY		142	VDCGGLLENEGTVYED-----	SGCRPLSCFME-----	PAPYTDSTGAAGGNN 186
DB		229	GLANGTVNDEPTTATDPEPSVVGVTSPAPLSTVTENPNNTSLPAPATPAEG--	-----	281
QY		187	CRFVESPSQD--QFLQARLRNPVRSLSQTPQNRPHGHSPELPE-----	-----	230
DB		282	---EESTGSGTQOLPA-----	---AAQAPDALPAGWEQRELPNGRVVYVDHNTKTTT 327	
QY		231	-----GYEQRITVQGVYFLHTQTGVSTWHDPRIERDLN-----	-----	264
DB		328	WERPLPGWEKRTDPRGRFYVDVHNTRTTTWQRPTAEVRYEQQVRSQRLQAGMHPFS 387		
QY		265	-----SVNCDLGLPLPGWEVRSVTSGRIFYVDVHNTRTTQFTDPR-----	-----	304
DB		388	QRFLYQSSASTDHPGLPLPGWEKQSD--NGRVVYVNHNTTRITQWEDPRPTQGMIOEPAL 446		
QY		305	-----LHHIMNHQCLKEPSQPLPLPSEGSLEDEELPAQSYERDLVQKLKVLK 352		
DB		447	PPGWEKMTYSEGVRIFYVDVHNTRITTKDPAPPGESGTKQSGPAYDSEFRWKVHQFRFLC 506		
QY		353	HELSQLQOAGACHRIEVSREIPEESYRQIMKMRKDKLKLKLVKFRGEGLDYGVGARE 412		
DB		507	HSNAL-----PSHWKISVSROTLEFSDSQQIMNMKPYDLRRRLYIIMRGERGLDYGGIARE 562		
QY		413	WLWLLCHEMLNPPYGLFOYSTDNIMQINPDSSINPDHLISYFHFVGRIMGLAVFHHYI 472		
DB		563	WFFLLSHEVLNPMYCLEFAYAKNNYCLQINPASSINPDHLITYFRFGRFTAMALYHGKFI 622		
QY		473	NGGFTVPFYKQLGKPIQLSDLESVDPELHKSJLVWILENDITPV--LDHTPCVVEHNAFGR 531		
DB		623	DTGFTTLPFYKRLMKRPTLXLSIDPEFYNSIVWIKENNLEBCLGLYFIQDMEILGV 682		
QY		532	LQHELKENGSRNVPTEENKEVYRLVYNNRPFMGIEBAQFLALQGNELIPIHLLPFDQ 591		
DB		683	TTHELKEGGSIRVTEENKEEYIMLLTDMFTFGVBEQTKAFDLGDNENAVLEWLYRFD 742		
QY		592	KELELIAGDKIDLNDKSNTRLKXHVADSNTRVRFQAVETFDSEERARILIQFVTGST 651		
DB		743	KELELMCGMQEIDMSDWQKSTIYRHYTKNSKQIQWFQVQVVKEMDNKRIILLOFVTGTC 802		
QY		652	RVPLQFKALQSGTGAAGRLFTIHLIDANTDNLPKAHTCFNRIDIPPEVSEKLYEKL 711		
DB		803	RLPVGGFAELIGSN---GPOKFCIDKVGKET--WLPSRHTCFNRLDLFPYKSYEQLEKLL 858		
QY		712	TAVEETCGFAVE 723		
DB		859	YAIETEFGQEE 870		

RESULT 15

NED4_MOUSE

ID

NED4_MOUSE

STANDARD;

PRT;

887 AA.

AC

P46935; O08758; O08GB3;

DT

01-NOV-1995 (Rel. 32, Created)

DT

10-OCT-2003 (Rel. 42, Last sequence update)

DT

10-OCT-2003 (Rel. 42, Last annotation update)

DE

Ubiquitin-protein ligase Nedd-4 (EC 6.3.2.-).

GN

NEDD4 OR NEDD-4 OR NEDD4A OR KIAA0093.

CC -!- FUNCTION: Involved in the embryonic development and differentiation of the central nervous system.
CC -!- SUBUNIT: Ubiquitin conjugation; third step.
CC -!- SUBUNIT: Binds SCNN1A, SCNN1B and SCNN1G (By similarity).
CC -!- TISSUE SPECIFICITY: Brain.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC -!- MISCELLANEOUS: A cysteine residue is required for ubiquitin-thioester formation.
CC -!- SIMILARITY: Contains 1 C2 domain.
CC -!- SIMILARITY: Contains 3 WW domains.
CC -!- SIMILARITY: Contains 1 HECT-type E3 ubiquitin-protein ligase domain.
CC -!- CAUTION: Ref.3 sequence differs from that shown due to a frameshift in position 12.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL; D85414; BAA12803.1; -
EMBL; U96635; AAB63360.1; ALT_FRAME.
EMBL; AK088620; BAC40458.1; -
EMBL; AK088767; BAC40558.1; -
EMBL; AK122203; BAC65485.1; -
HSP; Q13526; 1FIN.
MGD; MGI:97297; Nedd4.
GO; GO:0005829; C:cytosol; IDA.
GO; GO:000151; C:ubiquitin ligase complex; IPI.
GO; GO:000515; F:protein binding; IPI.
GO; GO:0004842; F:ubiquitin-protein ligase activity; IDA.
InterPro; IPR000008; C2_CaLB.
InterPro; IPR008973; C2_CaLB.
InterPro; IPR000569; HECT domain.
InterPro; IPR001202; WW_Rsp5_WWP.
Pfam; PF00168; C2; 1.
Pfam; PF00632; HECT; 1.
Pfam; PF00397; WW; 3.
SMART; SM00239; C2; 1.
SMART; SM00119; HECTC; 1.
SMART; SM00456; WW; 3.
PROSITE; PS00499; C2_DOMAIN_1; 1.
PROSITE; PS00004; C2_DOMAIN_2; 1.
PROSITE; PS50237; HECT; 1.
PROSITE; PS01159; WW_DOMAIN_1; 3.
PROSITE; PS50020; WW_DOMAIN_2; 3.
KW Ubl conjugation pathway; Ligase; Repeat.
FT DOMAIN 65 166 C2 DOMAIN.
FT DOMAIN 249 282 WW 1.
FT DOMAIN 405 438 WW 2.
FT DOMAIN 460 493 WW 3.
FT DOMAIN 552 887 HECT.
FT BINDING 854 854 UBIQUITIN (BY SIMILARITY).
SQ SEQUENCE 887 AA; 102705 MW; AE7DD3ED63986C50 CRC64;

Query Match 31.5%; Score 1225; DB 1; Length 887;
Best Local Similarity 33.6%; Pred. No. 1.3e-81;
Matches 291; Conservative 123; Mismatches 238; Indels 214; Gaps 23;

QY 7 IELTVLCAKNAKQFPFLPPAKIV-----VDGSGGCHSDTVKNTLDPKWCHY---- 58
DB 80 VRKRVIGIGLAKDILGASDPYKRVLYDPMGILTSVQTKLSLNPAWNEILFRV 139
QY 59 -----DLVVGKTD--SITISVWN-----HKIKHKQG 83
DB 140 LPQRHILFEVDENLRTRDFLGQVDVPLYPPTENPRMRYPTFKDFVLHPRSHKSRV 199
QY 84 AGFLGCVLLSNALRIKDT----- 103

Db 200 KGYL-----RLKMTYLPKNGSEDAQAEELEPGWVVLQDQDAATHLPHPPPEP 248
QY 104 ----GY-QRLDLC-----KLNPSDTDAVRGQVWVLSLQTRDRI 135
Db 249 SPLPGWEERQDVLGRYYVNHESRRQWKPSPDDDLTDEDNDMLQQAQRAFTTRQI 308
QY 136 GTGSSVVDRCGLLEN-----EGTVYEDSGPRPLSCFME-----EP 171
Db 309 SEDVDGPDNRESPENWEIVREDENTYSGOAVSPSPSCHIDVQTHLAEENFTRLAVCGNP 368
QY 172 A---PYTDTGAAAGG--NCRFVESPSQORQAQLRNPVDVRSGLQTPQNRPHGHQSP 226
Db 369 ATSPQVTSNHSNRGSLQTCIFEQFTLPVLLPT-----SS 405
QY 227 ELPEGYEORTTVQGVYFLHTQTGVSTWHD-----RIPRDL-NSVACDELGLPPPG 277
Db 406 GLPGWEEXQDGRSYVDHNSKTTTWSKPTMQDDPRSKIPAHLRGKTDSDNLGLPLPG 465
QY 278 WEVSTVSGRIYFVDHNRRTQTDPRLHHMHQCCKLPSQPLPLPSGSLDEELPA 337
Db 466 WEERTHTDGRVFFINHNKIKTQWEDPRLQNV-----AITGPAVP----- 504
QY 338 QRYERDLVOKLKLRLHLSLQPOAGHCRIEVSREEIFESYRQIMKWRPKD-LKXELMV 396
Db 505 --YSDDYKRYEYFFRKLKQTDIPNKFEMKLRANILEDSYRRIMGVGRADLLKARLWI 562
QY 397 XFRGEGLDYGVAREWLYLICHEMLAPYGLFOYSTDNIMQLINPDSSIT-NPDHLSYF 455
Db 563 EFDGKGLDYGVAREWFFLLISKEMFNYPYGLFEYSATDNYTLQINPNSGLCNEHLSYF 622
QY 456 HFVGRIMGLAVFHGYINGFTVPFYKOLLCKPIQLSLDSVDPDLHKSLLVILENDITP 515
Db 623 KFIGRVAGNAVTHGKLDGFFIRFYKMLQKLITLHDMESVDSEYSSLEWILENDPTE 682
QY 516 VLDHTFCVEHNAFRILQHELKPNRNVVTEENKGYRVLVYVNRVWFMGIEAQLQK 575
Db 683 -LDLRFIIDEELFGQTHQHELKGTGSEIVVTNKNKEVYLVIVQWRFVNRVQKQMAAFKE 741
QY 576 GENELIPOHLLKPPDOKELELIIGGLDKIDLNDKSNTELKHCVA-DSNIVRWFWQAVET 634
Db 742 GFELIPQDLIKIFDENELELLMCGLGVDVNDREHTKYKNGYSMMHQVHWFVKAWM 801
QY 635 FDEERRARLLQFVTGSTRVPLQGFKALQGSTGAAGPRLFTIHLIDANTDNLPAKHTCFNR 694
Db 802 MDSEKRIILQFVTGSTRVPMNGFAELVGSN---GPOSFTVEQW-GTPDKLPRAHTCFNR 857
QY 695 IDIPPEYSEKLYEKLLTAVEETCGF 720
Db 858 LDLPPYESFDELWDKQLQMAIENTQGF 883

Search completed: September 21, 2004, 07:44:16
Job time : 24.7111 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 21, 2004, 07:36:41 ; Search time 87.9789 Seconds
(without alignments)
2592.890 Million cell updates/sec

Title: US-10-009-945-2
Perfect score: 3884
Sequence: 1 GGSSKIRLTVCANLAKK.....EKLYKLTAVETCGFAVE 723

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPREMBL 25:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archaea:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3792.5	97.6	728	11 Q8K300	Q8K300 mus musculus
2	2206	56.8	1061	5 Q9V853	Q9V853 drosophila
3	1895	48.8	355	11 Q8BSC0	Q8BSC0 mus musculus
4	1559.5	40.2	295	4 Q8NDD8	Q8NDD8 homo sapien
5	1341	34.5	786	3 Q14326	Q14326 schizosacch
6	1340.5	34.5	288	4 Q96DE7	Q96DE7 homo sapien
7	1269	32.7	911	4 Q8N5A7	Q8N5A7 homo sapien
8	1247	32.1	955	4 Q96PU5	Q96PU5 homo sapien
9	1245	32.1	947	4 Q725F1	Q725F1 homo sapien
10	1245	32.1	975	4 Q725N3	Q725N3 homo sapien
11	1243	32.0	967	4 Q725F2	Q725F2 homo sapien
12	1241.5	32.0	820	4 Q9NT88	Q9NT88 homo sapien
13	1239	31.9	949	5 Q9Y0H4	Q9Y0H4 drosophila
14	1228.5	31.6	855	4 Q8WU09	Q8WU09 homo sapien
15	1228.5	31.6	858	4 Q8W58	Q8W58 homo sapien
16	1228.5	31.6	971	13 Q42573	Q42573 xenopus lae

ALIGNMENTS

RESULT 1
Q8K300 PRELIMINARY; PRT; 728 AA.
AC Q8K300;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DE Similar to E3 ubiquitin ligase SMURF1.
GN 4930431E10RIK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxID=10090;
[1]
RN SEQUENCE FROM N.A.
RA Strausberg R.;
RL Submitted (May-2002) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: CONTAINS 1 C2 DOMAIN.
DR EMBL; BC029097; AAH29097.1; -;
DR MGD; MGI:1923038; 4930431E10RIK.
DR GO; GO:0005622; C:intracellular; IEA.
DR GO; GO:0016874; F:ligase activity; IEA.
DR GO; GO:0004842; F:ubiquitin-protein ligase activity; IEA.
DR GO; GO:0006512; P:ubiquitin cycle; IEA.
DR InterPro; IPR000008; C2.
DR InterPro; IPR008973; C2 CaLB.
DR InterPro; IPR000569; HECT domain.
DR InterPro; IPR001202; WW_Rap5_WWP.
DR Pfam; PF00168; C2; 1.
DR Pfam; PF00632; HECT; 1.
DR Pfam; PF00397; WW; 2.
DR SMART; SM00239; C2; 1.
DR SMART; SM00119; HECTC; 1.
DR SMART; SM00456; WW; 2.
DR PROSITE; PS00499; C2 DOMAIN_1; 1.
DR PROSITE; PS50004; C2 DOMAIN_2; 1.
DR PROSITE; PS50237; HECT; 1.
DR PROSITE; PS01159; WW DOMAIN_1; 1.
DR PROSITE; PS50020; WW DOMAIN_2; 2.

O43165 homo sapien
Q9N2W4 homo sapien
Q8CF10 mus musculus
Q8BRT9 mus musculus
Q99PK2 mus musculus
Q95R64 drosophila
Q9N2Z7 caenorhabdi
Q95XU3 caenorhabdi
Q98TS4 xenopus lae
Q9CS63 mus musculus
Q8IQR6 drosophila
Q8TC08 drosophila
Q9VVI3 drosophila
Q80TB1 mus musculus
Q95TQ0 drosophila
Q9BKW4 caenorhabdi
Q9HCC7 schizosacch
Q9P2P5 homo sapien
Q8K4P8 mus musculus
Q8K325 drosophila
Q8BIA6 mus musculus
Q15029 homo sapien
Q8NG67 homo sapien
Q8VD14 mus musculus
Q7TMY8 mus musculus
Q921M5 mus musculus
Q8CFH2 mus musculus
Q9P4Z1 neuropsora

KW	Sequence	728 AA	83083 MW	C355291B9D8AD757 CRC64	
QY	2 GSSIKIRLTVLCAKNIKAKDPRLPDPRAKIVVDSGGCHSTDTVNTKLTDPKNNCHVDLY	61			
Db	10 GSSIKIRLTVLCAKNIKAKDPRLPDPRAKIVVDSGGCHSTDTVNTKLTDPKNNCHVDLY	69			
QY	62 VGTDSITISVNNKIKHKQAGAGLGCVRLLSNAISLKDGTQYRLDLCKLNPSDSDAV	121			
Db	70 VGTDSITISVNNKIKHKQAGAGLGCVRLLSNAISLKDGTQYRLDLCKLNPSDSDAV	129			
QY	122 RGOIVVSLQTRDRIGTGSVWDCRGLLENEGTVVSDSGRPLSCFEEPEAPYTDSTGAA	181			
Db	130 RGOIVVSLQTRDRIGTGSVWDCRGLLENEGTVVSDSGRPLSCFEEPEAPYTDSTGAA	189			
QY	182 AGGNCRCFVSPSQDORLQAOALRNPVGRSLQTPQNRPHGHQSPFLPEGYEQRTTVOGQ	241			
Db	190 AGGNCRCFVSPSQDORLQAOALRNPVGRSLQTPQNRPHGHQSPFLPEGYEQRTTVOGQ	249			
QY	242 VFLHTQTGVTWHPDRIPRDLNSVNCDELGPLPGWEVRSTVSGRIYFVDHNNRTTQFT	301			
Db	250 VFLHTQTGVTWHPDRIPRDLNSVNCDELGPLPGWEVRSTVSGRIYFVDHNNRTTQFT	309			
QY	302 DPLHHIMNHQCOLKEPSQPLPSPGSLDEBELPAQYERDLVOKLVRLHLSLQOQP	361			
Db	310 DPLHHIMNHQCOLKEPSQPLPSPGSLDEBELPAQYERDLVOKLVRLHLSLQOQP	369			
QY	362 AGHCRLEVSREIFESYRIQMKRPEKDLKRLMKVKGREGLOYGVAREWLYLLCHEM	421			
Db	370 AGHCRLEVSREIFESYRIQMKRPEKDLKRLMKVKGREGLOYGVAREWLYLLCHEM	429			
QY	422 LNPYGLFOYSDNTIYMLQINPDSSINPDHLSYHFVGRIMGLAVFRGHYINGGFTVPFY	481			
Db	430 LNPYGLFOYSDNTIYMLQINPDSSINPDHLSYHFVGRIMGLAVFRGHYINGGFTVPFY	489			
QY	482 KOLLGKPIQLSDLESVDPELHSLVWILENDITPVLDTHTFCVHNAGRILOHLEKNGR	541			
Db	490 KOLLGKPIQLSDLESVDPELHSLVWILENDITPVLDTHTFCVHNAGRILOHLEKNGR	549			
QY	542 NVPVTEENKKEYRVLVYVNRFMRGIEAQFLAQKGFNELIPQHLLKPFQKLELELIIGL	601			
Db	550 NVPVTEENKKEYRVLVYVNRFMRGIEAQFLAQKGFNELIPQHLLKPFQKLELELIIGL	609			
QY	602 DKIDLNDKSNTRLKHCVADSNVFWQAVETDEERRARLLQFVTGSTVPLQGFAL	661			
Db	610 DKIDLNDKSNTRLKHCVADSNVFWQAVETDEERRARLLQFVTGSTVPLQGFAL	669			
QY	662 QGSTGAAGRLFTIHLIDANTDNLPAKTCFNRIIDIPYSEYKLYEKLTAVEETCGFA	721			
Db	670 Q---GAAGRLFTIHLIDANTDNLPAKTCFNRIIDIPYSEYKLYEKLTAVEETCGFA	726			
QY	722 VE 723				
Db	727 VE 728				

RESULT 2

Q9V853

ID Q9V853 PRELIMINARY; PRT: 1061 AA.

AC Q9V853: Q9U3W2;

DT 01-WAY-2000 (TRENDELrel. 13, Created)

DT 01-OCT-2002 (TRENDELrel. 25, Last sequence update)

DT 01-OCT-2002 (TRENDELrel. 25, Last annotation update)

DE CG4943 protein (E3 ubiquitin ligase) (Ubiquitin-protein ligase) (Smad-

DE Ubiquitin E3 ligase (Smurfi1).

GN LACK OR SMURF1 OR CG4943.

OS Drosophila melanogaster (Fruit fly).

OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;

OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;

OC Ephydroidea; Drosophilidae; Drosophila.

OX NCBI_TaxID=7227;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=Berkley;

RX MEDLINE=20196006; PubMed=10731132;

RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scherz S.E., Li P.W., Hoskins R.A., Galle R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brannon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Heit G., Nelson C.R., Miklos G.L.G.,
 RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasly E.M.,
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
 RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,
 RA Burris K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cline S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
 RA Finkler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hoskins J., Houston K.A., Howland J.J., Wei M.-H., Ibegwam C.,
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Laslo P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Pachter J.M.,
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svirska R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissenbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zhang L.,
 RA Zheng X.H., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.N., Rubin G.M., Venter J.C.;
 RT "The genome sequence of *Drosophila melanogaster*;"
 RL Science 287:2185-2195(2000).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Celniker S.E., Adams M.D., Krommiller B., Wan K.H., Holt R.A.,
 RA Evans C.A., Gocayne J.D., Amanatides P.G., Brandon R.C., Rogers Y.,
 RA Benson J., An H., Baldwin D., Banon J., Beeson K.Y., Busam D.A.,
 RA Carlson J.W., Center A., Champe M., Davenport L.B., Dietz S.M.,
 RA Dodson K., Dorsett V., Doup L.E., Doyle C., Dresnek D., Farfan D.,
 RA Ferreira S., Frise E., Galle R.F., Garg N.S., George R.A.,
 RA Gonzalez M., Houck J., Hoskins R.A., Hostin D., Howland T.J.,
 RA Ibegwam C., Jalali M., Kruse D., Li P., Mattei B., Moshrefi A.,
 RA McIntosh T.C., Moy M., Murphy B., Nelson C., Nelson K.A., Nunco J.,
 RA Pachter J., Parages V., Park S., Patel S., Pfeiffer B.,
 RA Phouanavong S., Pittman G.S., Puri V., Richards S., Scheeler F.,
 RA Stapleton M., Strong R., Svirska R., Tector C., Tyler D.,
 RA Williams S.M., Zaveri J.S., Smith H.O., Venter J.C., Rubin G.M.;
 RT "Sequencing of *Drosophila melanogaster* genome;"
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RA Misra S., Crosby M.A., Matthews B.B., Bayraktaroglu L., Campbell K.,
 RA Hradecky P., Huang Y., Kaminker J.S., Prochuk S.E., Smith C.D.,
 RA Tupy J.L., Bergman C., Berman B., Carlson J.W., Celniker S.E.,
 RA Clamp M., Drysdale R., Emmert D., Frise E., de Grey A., Harris N.,
 RA Krommiller B., Marshall B., Millburn G., Richter J., Russo S.,
 RA Searle S.M.J., Smith E., Shu S., Smutniak F., Whitfield E.,
 RA Ashburner M., Gelbart W.M., Rubin G.M., Mungall C.J., Lewis S.E.;
 RT "Annotation of *Drosophila melanogaster* genome;"
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 RN [4]

```

SEQUENCE FROM N.A.
RA Adams M.D., Ceiniker S.E., Gibbs R.A., Rubin G.M., Venter C.J.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [5]
RP SEQUENCE FROM N.A.
RA FlyBase;
RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
RN [6]
RP SEQUENCE FROM N.A.
RA Laurencon A., Hawley S.;
RL "Molecular cloning of a type E3 Ubiquitin ligase.";
RT Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
RN [7]
RP SEQUENCE FROM N.A.
RA Podos S.D., Hanson K.K., Wang Y.-C., Ferguson E.L.;
RL "The Dsmurf ubiquitin-protein ligase restricts BMP signaling spatially
and temporally during Drosophila development.";
RN [8]
RP SEQUENCE FROM N.A.
RA Liang Y.-Y., Lin X., Feng X.-H.;
RL "dSmurf1, a Smad-ubiquitin E3 ligase, specifically targets dpp-
activated Mad-protein to the EMBL/GenBank/DBJ databases.
CC -I- SIMILARITY: CONTAINS 1 C2 DOMAIN.
DR EMBL; AE003802; RAF57824.3; -
DR EMBL; AF216521; RAF21125.1; -
DR EMBL; AF416571; RAL09691.1; -
DR EMBL; AF464851; AMO9646.1; -
DR HSSP; Q13526; IPIN.
DR FlyBase; FBgn0029006; lack.
DR GO; GO:0005622; C:intracellular; IEA.
DR GO; GO:0016874; F:ligase activity; IEA.
DR GO; GO:0004842; F:ubiquitin-protein ligase activity; IEA.
DR GO; GO:006512; P:ubiquitin cycle; IEA.
DR InterPro; IPR000008; C2.
DR InterPro; IPR008973; C2_CalB.
DR InterPro; IPR000569; HECT domain.
DR InterPro; IPR001202; WW_Rsp5_WWP.
DR Pfam; PF00168; C2; 1.
DR Pfam; PF00632; HECT; 1.
DR Pfam; PF00397; WM; 3.
DR SMART; SM00239; C2; 1.
DR SMART; SM00456; WW; 3.
DR PROSITE; PS00499; C2_DOMAIN_1; 1.
DR PROSITE; PS00004; C2_DOMAIN_2; 1.
DR PROSITE; PS0237; HECT; 1.
DR PROSITE; PS01159; WW_DOMAIN_1; 2.
DR PROSITE; PS00020; WW_DOMAIN_2; 3.
KW Ligase.
SQ SEQUENCE 1061 AA; 115675 MW; 6BBC550F5129163 CRC64;

Query Match          56.8%; Score 2206; DB 5; Length 1061;
Best Local Similarity 43.0%; Pred. No. 5.3e-172;
Matches 454; Conservative 108; Mismatches 148; Indels 346; Gaps 14

QY      6 KRLTLVLCAKNAIKDFFRLDDPPAKIVVDGSGCHSTDVKNLTDPKNNQHLYLVGKT 65
Db      14 KYRITLCARNLARKDLFRLEDPFAKVQDGTGVGSTEISKSLSLPKWNAYDLFLGIG 73
QY      66 DSITTSVNHHKHKKOCAGFLGCVRLLSNAIRLKDTGYQRLLDKLNPSPTDAVRGT 125
Db      74 DAIIITVWNRKIHK--GSGLFGCVIRPAFNIQSLKGAGFORLDLKLSPDDDELVRGI 131
QY      126 VVSLQTRDIRGTG-----GSVPDCRGILLE-----GTVV----- 155
Db      132 ITSLSKDGSPGNPLAIVGPSGDVRGPSDDSDSLPGWEERTDNGRVVYNHATK 191
QY      156 ----- 155
Db      192 STQWRDPRPGVGVSSHATSPOQRHNTHNGSGDRQAAPGRTRSTTCTNLMMNGHSRD 251
QY      156 -----EDSGFRFLSCFMEEPAFYTDSTGAAGGNCRFVE---- 191

```

RA The FANTOM Consortium,
 RA The RIKEN Genome Exploration Research Group Phase I & II Team;
 RT "Analysis of the mouse transcriptome based on functional annotation of
 RT 60,770 full-length cDNAs";
 RL Nature 420:563-573 (2002).
 DR EMBL; AK034736; BAC28813.1; -.
 DR MGI; 1923038; 4930431B10Rik.
 DR GO; GO:0005622; C:intracellular; IEA.
 DR GO; GO:0004842; F:ubiquitin-protein ligase activity; IEA.
 DR GO; GO:0006512; P:ubiquitin cycle; IEA.
 DR InterPro; IPR000569; HECT_domain.
 DR Pfam; PF00632; HECT; 1.
 DR SMART; SM00119; HECTC; 1.
 DR PROSITE; PS50237; HECT; 1.
 KW Hypothetical protein.
 FT NON TER 1
 SQ SEQUENCE 355 AA; 41315 MW; 93B39B4C82F86DD1 CRC64;
 Query Match 48.8%; Score 1895; DB 11; Length 355;
 Best Local Similarity 99.7%; Pred.No. 3.7e-147; Indels 0; Gaps 0;
 Matches 354; Conservative 0; Mismatches 1;
 QY 369 VSREIEFEESVQIMKMRPDKLKKLMLVKFRGEGLDYGGVAREWLYLLCHEMLNPFYGL 428
 Db 1 VSREIEFEESVQIMKMRPDKLKKLMLVKFRGEGLDYGGVAREWLYLLCHEMLNPFYGL 60
 QY 429 FOYSTDNIMQLINPDSSINPDHLSYFHFVGRIMGLAVFHGHYINGGFTVPFYKOLLGKP 488
 Db 61 FOYSTDNIMQLINPDSSINPDHLSYFHFVGRIMGLAVFHGHYINGGFTVPFYKOLLGKP 120
 QY 489 IQLSDLESVDPELHKSLVWILENDITPVLDHTFCVEHNAFGRILQHELKPKNGRNPVT 548
 Db 121 IQLSDLESVDPELHKSLVWILENDITPVLDHTFCVEHNAFGRILQHELKPKNGRNPVT 180
 QY 549 NKKEYVRLYVNWRFMRGIEAQFLALQKGFNELIPQHLLKPFDPCKEELIIGGLDKIDLND 608
 Db 181 NKKEYVRLYVNWRFMRGIEAQFLALQKGFNELIPQHLLKPFDPCKEELIIGGLDKIDLND 240
 QY 609 WKSNTLKHCVADSNIRVWFQAVETFEERRARLLQFVTGSTRVPLQGFKALQSGTGAA 668
 Db 241 WKSNTLKHCVADSNIRVWFQAVETFEERRARLLQFVTGSTRVPLQGFKALQSGTGAA 300
 QY 669 GPRFLTTHLIDANTDNLPKHTCFNRIDIPPYSEYKLYEKLITAVEETCGFAVE 723
 Db 301 GPRFLTTHLIDANTDNLPKHTCFNRIDIPPYSEYKLYEKLITAVEETCGFAVE 355
 RESULT 4
 QNDD8
 ID Q8NDD8 PRELIMINARY; PRT; 295 AA.
 AC Q8NDD8
 DT 01-OCT-2002 (TRENBLrel. 22, Created)
 DT 01-OCT-2002 (TRENBLrel. 22, Last sequence update)
 DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)
 DE Hypothetical protein (Fragment).
 GN DKFZP564H223.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RA Wambutt R., Heubner D., Mewes H.W., Gassenhuber J., Wiemann S.;
 RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AL834242; CAD38919.1; -.
 DR GO; GO:0005622; C:intracellular; IEA.
 DR GO; GO:0004842; F:ubiquitin-protein ligase activity; IEA.
 DR GO; GO:0006512; P:ubiquitin cycle; IEA.
 DR InterPro; IPR000569; HECT_domain.
 DR Pfam; PF00632; HECT; 1.
 DR SMART; SM00119; HECTC; 1.
 DR PROSITE; PS50237; HECT; 1.

KW Hypothetical protein.
 FT NON TER 1
 SQ SEQUENCE 295 AA; 34030 MW; 0507325127A943EA CRC64;
 Query Match 40.2%; Score 1559.5; DB 4; Length 295;
 Best Local Similarity 98.7%; Pred.No. 1.1e-119; Indels 3; Gaps 1;
 Matches 293; Conservative 1; Mismatches 0;
 QY 427 GLFOYSTDNIMQLINPDSSINPDHLSYFHFVGRIMGLAVFHGHYINGGFTVPFYKOLLG 486
 Db 2 GLFOYSTDNIMQLINPDSSINPDHLSYFHFVGRIMGLAVFHGHYINGGFTVPFYKOLLG 61
 QY 487 KPIQLSDLESVDPELHKSLVWILENDITPVLDHTFCVEHNAFGRILQHELKPKNGRNPVT 546
 Db 62 KPIQLSDLESVDPELHKSLVWILENDITPVLDHTFCVEHNAFGRILQHELKPKNGRNPVT 121
 QY 547 BENKKEYVRLYVNWRFMRGIEAQFLALQKGFNELIPQHLLKPFDPCKEELIIGGLDKIDL 606
 Db 122 BENKKEYVRLYVNWRFMRGIEAQFLALQKGFNELIPQHLLKPFDPCKEELIIGGLDKIDL 181
 QY 607 NDWKSNTLKHCVADSNIRVWFQAVETFEERRARLLQFVTGSTRVPLQGFKALQSGTG 666
 Db 182 NDWKSNTLKHCVADSNIRVWFQAVETFEERRARLLQFVTGSTRVPLQGFKALQSGTG 238
 QY 667 AAGPRFLTTHLIDANTDNLPKHTCFNRIDIPPYSEYKLYEKLITAVEETCGFAVE 723
 Db 239 AAGPRFLTTHLIDANTDNLPKHTCFNRIDIPPYSEYKLYEKLITAVEETCGFAVE 295
 RESULT 5
 OI4326
 ID OI4326 PRELIMINARY; PRT; 786 AA.
 AC OI4326
 DT 01-JAN-1999 (TRENBLrel. 09, Created)
 DT 01-JAN-1999 (TRENBLrel. 09, Last sequence update)
 DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)
 DE Ubiquitin--protein ligase.
 GN SPBC1689.11C.
 OS Schizosaccharomyces pombe (Fission yeast).
 OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
 OC Schizosaccharomycetales; Schizosaccharomycetaceae;
 OC Schizosaccharomycetes.
 OX NCBI_TaxID=4896;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=972h-;
 RA Volckaert G., Wood V., Rajandream M.A., Barrell B.G.;
 RL Submitted (AUG-1997) to the EMBL/GenBank/DBJ databases.
 CC -1- SIMILARITY: CONTAINS 1 C2 DOMAIN.
 DR EMBL; Z99759; CAB16903.1; -.
 DR FIR; T39585; T39585.
 DR HSSP; Q13526; 1PIN.
 DR GeneDB; SPombe; SPBC169.11c; -.
 DR GO; GO:0005622; C:intracellular; IEA.
 DR GO; GO:0016874; F:ligase activity; IEA.
 DR GO; GO:0004842; F:ubiquitin-protein ligase activity; IEA.
 DR GO; GO:0006512; P:ubiquitin cycle; IEA.
 DR InterPro; IPR000008; C2.
 DR InterPro; IPR008973; C2_CalB.
 DR InterPro; IPR000569; HECT_domain.
 DR InterPro; IPR002349; WW.
 DR InterPro; IPR001202; WW_Rsp5_WWP.
 DR Pfam; PF00168; C2; 1.
 DR Pfam; PF00632; HECT; 1.
 DR Pfam; PF00397; WW; 3.
 DR PRINTS; PR00403; WWDOMAIN.
 DR SMART; SM00239; C2; 1.
 DR SMART; SM00119; HECTC; 1.
 DR SMART; SM00456; WW; 3.
 DR PROSITE; PS00499; C2_DOMAIN_1; 1.
 DR PROSITE; PS50004; C2_DOMAIN_2; 1.
 DR PROSITE; PS50237; HECT; 1.
 DR PROSITE; PS01159; WW_DOMAIN_1; 3.

OS	Homo sapiens (Human).
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX	NCBI_TaxID=9606;
RN	[1]
RP	SEQUENCE FROM N.A.
RC	TISSUE=Lung;
RA	Strausberg R.;
RL	Submitted (JUN-2001) to the EMBL/GenBank/DDBJ databases.
DR	EMBL; BC009527; AAH09527.1; -
DR	GO; GO:0005622; C:intracellular; IEA.
DR	GO; GO:0016874; F:ligase activity; IEA.
DR	GO; GO:0004842; F:ubiquitin-protein ligase activity; IEA.
DR	GO; GO:0006512; P:ubiquitin cycle; IEA.
DR	InterPro; IPR000569; HECT_domain.
DR	Fam; PF00632; HECT; 1.
DR	SMART; SM00119; HECTC; 1.
DR	PROSITE; PS0237; HECT; 1.
KW	Ligase.
FT	NON TER
SQ	SEQUENCE 288 AA; 33255 MW; PE2B43E300BE6537 CRC64;
Query Match 34.5%; Score 1340.5; DB 4; Length 288;	
Best Local Similarity 84.8%; Pred.No.1e-101;	
Matches 246; Conservative 23; Mismatches 18; Indels 3; Gaps 1;	
QY	434 DNIYMLQINPDSINSNPDLHSYEHFVGSIMGLAVFHGHVINGGVVPFYKOLLGPIQLSD 493
Db	2 DDITTYQINPDSAVNEHLSYFFHVGRIOMGVFHHGYIDGGFTLPFYKOLLGKSITLDD 61
QY	494 LESVDPELHKSLVMILENDITPVLDHTFCVEHNAFGRILQHCLKPNGRNVPTVEENKEY 553
Db	62 MELVDPDLNSLVILENDITGVLDHTFCVEHNAGBELIQHELKPKNGKSPVNEENKEY 121
QY	554 VRLIYVWRFRNGTEAQFLALOKGFNELIPOLHLKPDPDKOLELIIGGLDKIDLMDWSNT 613
Db	122 VRLIYVWRFLRGTEAQFLALOKGFNEVIPOLHLKTDFDEKELELICGLGKIDVNDWKVNT 181
QY	614 RLKHCVADSNIVWFVQVAETPDERRARLLQVTGTSTRVPLQGFKALOGSTGAAGRPLF 673
Db	182 RLKHCTPDSNIVKFVKNAVFEDERRARLLQVTGSSRVPLOGFKALO--GAAGRPLF 238
QY	674 THILIANTNLKPAHTCFNRIDIPPYESYEKLYEKLLTAVERTCGFAVE 723
Db	239 TIHQIDACTNNLKPAHTCFNRIDIPPYESYEKLYEKLLTAIERTCGFAVE 288
RESULT 7	
Q8NSA7	PRELIMINARY; PRT; 911 AA.
ID	Q8NSA7
AC	Q8NSA7;
DT	01-OCT-2002 (TrEMBLrel. 22, Created)
DT	01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT	01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE	Similar to neural cell expressed, developmentally down-regulated
DE	4-like.
OS	Homo sapiens (Human).
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX	NCBI_TaxID=9606;
RN	[1]
RP	SEQUENCE FROM N.A.
RC	TISSUE=Melanoma;
RA	Strausberg R.;
RL	Submitted (JUN-2002) to the EMBL/GenBank/DDBJ databases.
CC	-1- SIMILARITY: CONTAINS 1 C2 DOMAIN.
DR	EMBL; BC032597; AAH32597.1; -
DR	GO; GO:0005622; C:intracellular; IEA.
DR	GO; GO:0004842; F:ubiquitin-protein ligase activity; IEA.
DR	GO; GO:0006512; P:ubiquitin cycle; IEA.
DR	InterPro; IPR000008; C2.
DR	InterPro; IPR0008973; C2_CalB.
DR	InterPro; IPR000569; HECT domain.

Query Match	32.1%	Score 1247;	DB 4;	Length 955;	
Best Local Similarity	31.9%;	Pred. No. 3.1e-93;			
Matches 310;	Conservative 125;	Mismatches 246;	Indels 292;	Gaps 24;	
QY	2	GSSIKIRLTVLCANLAKDQFRLDPFAKI	---VVDGSGQCH---STDVKNLTDPKWNQ	56	
DB	17	GESRILRVKVVSGIDLAKDIFGASDPVKLSLVADENRELALVQTKIKTLINPKNE	76		
QY	57	HYDLVCKTD-SITISVWNHKKHKQAGFLGCVRLLSNAI	---155		
DB	77	EFYFRVNPNSHRLLEFVFDENRLTRDD	---FLGQVDVPLSHLPDFTMERPTYFKDFLL	133	
QY	98	-----SRLKDGTGQRLDCLKNPSTDVARGQIVVSLQTRDRTIGTGGVVDCRG	-----146		
DB	134	RPRSHKSRVK--GFLRLKMAYPKNG	-----GQDEENSQDQDDMEHCWVVDSDNSASQH	186	
QY	147	-----LLENGTIV	-----155		
DB	187	QBELPPPLPPGWEKVDNLGRITYVNNHNRRTQWHRPSLMDVSSDNNIRQINQEAH	246		
QY	156	-----EDSGP	-----GRPLSCFMEEPAPTYDSTGAA	-----AGGNCREVES	192
DB	247	RFRSRRRHISEDLEPEPESEGDPPEPWETISEVNIAGDSLGLALPPPPASPGSRTSPQE	306		
QY	193	PSOD	-----QRLCAQRLN	-----206	
DB	307	LSELSRRLQITPDSNGEQFSSLIQRPESRLRSCSVTDVABQGLHPLPPSVAYVHTTPG	366		
QY	207	-----	-----PDVR	-----210	
DB	367	LPGWEERKDAKGRITYVNNHNRRTTWTREIMQLAEDGASGSATSNHNLIEQIRAPRS	426		
QY	211	-----	-----GSLQTPNPHGHQSP	-----LPEGYEQR	235
DB	427	LSSEPTVLSAPLEGAKDSPVRAVKDTLSNPQSPQSPYNSPKQHKVQSTQFPPGWEHR	486		
QY	236	TTVQGVYFHTQTGVSTWHDRIIP	-----RDLSVNCDELGLPPLPGWEYRSTVSGRIYF	290	
DB	487	IAPNGRPFFTDHNTKITTTWEDPRLKPPVHMRKSTSLNPNDLGPLPGWEERIHLDGRTFY	546		
QY	291	VDHNNRTQFTDRLHHNNHQCLKEPSQPLPLPSEGSLEDELPACRYERDLVQKLKV	350		
DB	547	IDNSKITQWEDPRLQV	-----PAITG	-----FAPVYSREFKOKDY	583
QY	351	LRHELSQLQPAQCHRIEVSREBIFBESYRQIKMK-RPKDLKKRLMVKFRGEBGLDYG	409		
DB	584	FRKLLAKPADIPNRFEMLKLRHNIPEBSYRIRMSVKRPDLKARLWTEPSEKGLDYG	643		
QY	410	AREWLYLICHEMLNPYYGLFCYTDNIYMLQIPDSSI-NPDHLSYFHFVGRIMGLAVPH	468		
DB	644	ARWFFLLKEMFNYYGLFYEYATNTYLIQINPNSGLCNEDHLSYFTFTFGRVAGLAVPH	703		
QY	469	GHYINGFTVPFYKQLGKPIQLSDLESVDPELHLSLWILENDITPVLDPHTFCVSHNAF	528		
DB	704	GKLLDGGFFIRFPYKMLMGKQITLNDMESVDSEYNSLKWILENDPTE-LDLMFCDIBENF	762		
QY	529	GRILQHELKFNRMVPTENKKEYVRLVYVWFMRGIEAOFALOKGFNELTPHLLKP	588		
DB	763	QGYQYVDLKPNGSEIWTNENKREYIDLVIQWRFVNRVQKMAFLEGFTPELLFIDLIKI	822		
QY	589	FDQKLELIGGLDKIDLNDKWSNTRLKHCVADSN-IVRMFWQAVETFDSEERARLLQFV	647		
DB	823	FDSENELELLMCGLDGVDVNDWRQHSYKNGCYCNHPVIQWFWKAVLLMDAEKRIALLQFV	882		
QY	648	TGSTRVPLQGFKALQGSTGAAGPRLFTIHLIDANTDNLPKAHTCFNRIDIPPYVESYKLY	707		
DB	883	TGTSRVPMNGFAELYSN---GPQLFTIEQW-GSPKLPRAHTCFNRLLDLPPTYETEDLR	938		
QY	708	EKLLTAVETCGF	720		
DB	939	EKLLMAVENAQGF	951		

```

RESULT 9
Q725F1
ID Q725F1 PRELIMINARY; PRT; 947 AA.
AC Q725F1;
DT 01-OCT-2003 (TEMBLrel. 25, Created)
DT 01-OCT-2003 (TEMBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TEMBLrel. 25, Last annotation update)
DE Ubiquitin ligase NEDD4h.
GN NEDD4L.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Qi H., Labrie C.;
RT "New splicing isoform of human Nedda4-2.";
RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY112985; AAM76730.1; -.
KW Ligase.
SQ
SEQUENCE 947 AA; 109429 MW; 95CE0FDEDCB96639 CRC64;
Query Match 32.1%; Score 1245; DB 4; Length 947;
Best Local Similarity 31.9%; Pred. No. 4; Se-93;
Matches 310; Conservative 125; Mismatches 246; Indels 292; Gaps 24;
QY 2 GSSIKIRLVLCAKNIAXKDFELPLPEAKI---VVDGSGQCH--STDIVKNTLDPKWQ 56
DB 9 GESRIIRKVVSGIDIAKDIQFASDPTVKLSLYVPEENRELALVQTKIKTLNPKWNE 68
QY 57 HYDLVYKGTD-SITISVNNHKIHKKGAGFLGCVLLSNAI----- 97
DB 69 EFYFRVNPNSHLLFVPEENLRTRD---FLGQVDVPLSHLPTEDPTMERPYTKDFLL 125
QY 98 -----SRLKDTGYQRLDLCKLNPSTDVAVGQIVVSLQTRDRIGTGSGVWDCRG--- 146
DB 126 RPASHKSRVK--GFLRKWAYMPKNG-----GQDEENSQORDDMEHGWVWDSNDSQ 178
QY 147 -----LLENEGTVY----- 155
DB 179 QEELPPPLPGWEEKVDNLGRITYVYVNNHNRITQWHRPSLMDVSSSDNNIRQINQEA 238
QY 156 -----EDSGP-----GRPLSCFMEEPAPYTDSTGAA-----AGGNCRFV 192
DB 239 RFRSRRHISEDLEPEPSEGQDVPEWETISEVNIAGSLGLALPPPPASPGSRTPQ 298
QY 193 PSQD-----ORLQACLRN----- 206
DB 299 LSEELSRRIQITPDSNGEQFSLIQEPEPSRLRSCVTDVAEQHLPSPVAVYVHTTG 358
QY 207 -----PDVR----- 210
DB 359 LPFGWBEERKDAKRTYVYVNNHNRITTTWTPIMQLAEDGASGATNSNNHLIEPQIR 418
QY 211 -----GSLQTPQNRHGHOSPE-----LPGYEOR 235
DB 419 LSSPTVTLSPGLEAGKDSVPRRAVKDTLSNPQSPSPYNSPKQHKVTQSFLLPQW 478
QY 236 TTVOGWYFLHTQTGVSTWHDPRIP-----RDLSNVNCDLGLPLPGWEVRSTVSGRI 290
DB 479 IAPNGRPPIIDHTKTTTWEDPLKFPVHMRSKTSLNPNDLGLPLPGWEERHLDGRTY 538
QY 291 VDHNNRTQTFTPRLHHIMHOCQKEFSQPLPSEGSLEDEELPAQRYERDLVQKLKV 350
DB 539 IDHNSKITQWEDPRQN-----FAITG-----PAVYSEEFKQKYD 575
QY 351 LRHELSLQOPQACHRCIEVRSEBEIPEESYRQIMKX-RPKXDLKKRLMKVPRGEGLDY 409
DB 576 FRKKLKKPADINRFEMKLRHNNIFEESYRIMSVKRPDVLKARLWIFESKGLDYG 635
QY 410 AREWYLLCHEMLNPIYGLFQYSTONTIYMLQINPDSSI-NPDHLSYFVHVGINGLA 468
DB 636 AREWFFLLSKEMFNPIYGLFEYSACNTYLTQINPNSGLCNEPHLSYFFPIGRVAG 695

```


PALAZZOLO M., PITMAN G.S., PAN S., POLLARD J.J., PURI V.V., REESE M.G.,

RRA
RRA
RRA
RRA
RRA
RRA
RRA
RRA
RRA
RRA
RRA
RRA
RRA
RRT
RRN
RRP

Shue B.C., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
Spier E., Specting A.C., Stapleton M., Strong R., Sun E.,
Swirskas R., Turner C., Turner R., Venter E., Wang A.H., Wang X.,
Wang Z.-Y., Wassman D.A., Weinstein G.M., Weissbach J.,
Williams S.W., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao O., Zheng L.,
Zheng X.H., Zhong F.N., Zhou W., Zhou X., Zhu S., Zhu X., Smith H.O.,
Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.:
"The genome sequence of *Drosophila melanogaster*";
science 287:2185-2195(2000).

[2]

SEQUENCE FROM N.A.
Cornell M., Evans D.A.P., Mann R., Postier M., Flaszka M.,
Monthonat M., Artavanis-Tsakonas S., Baron M.;
"The *Drosophila melanogaster* Suppressor of deltex gene, a regulator of
the Notch receptor signalling pathway, is an E3 class ubiquitin
ligase";
Genetics 152:0-0(1999).
-1- SIMILARITY: CONTAINS 1 C2 DOMAIN.
EMBL; AE003584; AF51312.1; -.
EMBL; AF152865; AAC39975.1; -.
HSSP; Q13526; IPIN.
FlyBase; FBgn0003557; Su(dx).
GO; GO:0007219; P:N signaling pathway; IGI.
GO; GO:0008587; P:wing margin morphogenesis; IGI.
GO; GO:0008586; P:wing vein morphogenesis; IGI.
InterPro; IPR000008; C2.
InterPro; IPR0008973; C2_CalB.
InterPro; IPR000569; HECT_domain.
InterPro; IPR001202; WW_Rsp5_WWP.
Pfam; PF00168; C2; 1.
Pfam; PF00632; HECT; 1.
Pfam; PF00397; WW; 4.
SMART; SM00239; C2; 1.
SMART; SM00119; HECT; 1.
SMART; SM00456; WW; 3.
PROSITE; PS50004; C2_DOMAIN_2; 1.
PROSITE; PS50237; HECT; 1.
PROSITE; PS01159; WW_DOMAIN_1; 3.
PROSITE; PS50020; WW_DOMAIN_2; 2.

SQ

Query Match 31.9%; Score 1239; DB 5; Length 949;
Best Local Similarity 32.3%; Pred. No. 1.4e-92;
Matches 301; Conservative 133; Mismatches 251; Indels 246; Gaps 27;

Qy 8 RLTVLCAKNAKKDFRLDPDPFAKIIVDGSOGCHSTDKVNTLPKNOHVDLVVGTKDS 67
Ddb :|::||::||::||::||::||::||::||::||::||::||::||::||::||:
50 QLSVTTEASLRNGFLKPVELLIISK-RKTDLVKNLYPKWNEETVLITNST 108
Qy :|::||::||::||::||::||::||::||::||::||::||::||::||::||:
68 ITTISWNHKHIKK-QAGFLGCVRLL-----SNATSLRKUDTG- 105
Ddb :|::||::||::||::||::||::||::||::||::||::||::||::||::||:
109 LHFKVLDDHSFRKDMLGERIINLAHLQHNGRCFELELTIDLFTSKSDNRQTSGEL 168
Qy :|::||::||::||::||::||::||::||::||::||::||::||::||::||:
106 -----QRDLCKL-----NPS-DTDAVRGQ---LVSLSQTRDI-- 135
Ddb :|::||::||::||::||::||::||::||::||::||::||::||::||::||:
169 VALINGLKLDMSKLQIPVAGOQNPNPVQAVNPFSWDSAAGSCMTYGGVARMRLRS 228
Qy :|::||::||::||::||::||::||::||::||::||::||::||::||::||:
136 ----GTGGSVVDCRGLENEGTYEDSGPGRPL-----SCFWEEPAPYDTDGAA--- 181
Ddb :|::||::||::||::||::||::||::||::||::||::||::||::||::||:
229 SSGNSNGC---EYTSPLPGCGGDHRSTRQPFWEQQQQQSQNQQLRWYNGSAAVPQ 285
Qy :|::||::||::||::||::||::||::||::||::||::||::||::||::||:
182 -----ACGNHC-----RFVESP---- 193
Ddb :|::||::||::||::||::||::||::||::||::||::||::||::||::||:
286 TAPYPQPPAPALARPLTOYGALPENTOPAAAYLPAGGGAAGVPGVAGGPPIEGQGVL 345
Qy :|::||::||::||::||::||::||::||::||::||::||::||::||::||:
194 ----SODORLQAQ-----RLR-----NP 207
Ddb :|::||::||::||::||::||::||::||::||::||::||::||::||::||:
346 PVSGSTDPLQOTPADDEPLPWAGEIRLDQYGRYYVDHTYSTYWEKPTLPFGWEIRK 405
Qy :|::||::||::||::||::||::||::||::||::||::||::||::||::||:
208 DVRGSL-----QTFQNRPHG-----HOSP----- 226

```
Db 406 DGRGRVYVDHNRKTTWQRPNSRLMHFQHQWQRAHVVSQGNRYLYSQQQOQPTAVT 465
Qy 227 -----ELPEGVEQRTTVGGYVFLHTQGVSTWHDPRIP-RDLNSVNCDELGP 273
Db 466 AQVQDDEDALQPLDGEKKIQSNRVIFVNHKRTTQWEDPRTOGQEVSLIN---EGP 522
Qy 274 LPPGWEVSTVSGRIYFVDHNNRTTQFTDPRLUHHIMHQCQLKEPSQPLPLPSSGLSDE 333
Db 523 LPPGWEIRVTAAGERFFVDHNNRTTFFEDPR-----PGAP-----XGAGVY 564
Qy 334 ELPAQRYERDLVOKLVRLHLSLQOPQAGHCRIEVSREEIFEEYSYROIIMKRPKDLKKR 393
Db 565 GVP-RAYERSFRWKLSPRY-LCQSNALPSHKTITVTRQTLFEDSYHQIMRLPAYELRR 622
Qy 394 LNVKFRGEGLDYGGVAREWLYLLCHEMLNPYYGLFOYSTNIYMLQINPDSSINPDHLS 453
Db 623 LVIIIFRGEGLDYGGVREWFELLSHEVLNPMYCLFEYANKUNYSLQINPASYNPDHLQ 682
Qy 454 YPHFVGRINGLAVFHGYINGFTVPFYKQLLGKPIQLSDLESVDPELHKSILVILENDI 513
Db 683 YPKFIGRFAMALYHGRFTYSGFTMPFYKRMNLKLTIDTIDTIDTIDTIDTIDTIDTID 742
Qy 514 TPV-LDHTFCVHNHAFGRILQHELKPNGRNVPVTEENKKEYVRLYVNRFRMGTEAQLA 572
Db 743 DCGGLWLFVSDFEVLGQIIHHELKENGKEKERTVEENKEEYITLMTWRMTRGIEQOQKT 802
Qy 573 LQKGNELIPOHLKPPDOKELELIIGLDKIDLNDWKSNTLRKHCVDNSIVRWFQAV 632
Db 803 FLEGNEVPLEWLYDFDERELICGMDQDVEDQWRNTIYRHNRSKQVVMFQFV 862
Qy 633 EYFDERRARLQFVTSRVPVLOGFKALQGSTGAAGRLFTIHLIDANTDNLKPAHCTCF 692
Db 863 RTDNEKRARLQFVGTGRVGVGFAELMGSN---GPQFCEIKVGKET-WLPSRSTCF 918
Qy 693 NRIDIPPEYSEKYLEKLITAVEETCGFAVE 723
Db 919 NRIIDLPYKSYDQVLEKLTFAETEETGFCQE 949

RESULT 14
Q8WUJ9
ID Q8WUJ9 PRELIMINARY; PRT; 855 AA.
AC Q8WUJ9
DT 01-WAR-2002 (TrEMBLrel. 20, Created)
DT 01-WAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]_TaxID=9606;
RP SEQUENCE FROM N.A.
RC TISSUE=Uterus;
RA Strausberg R.;
RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC019345; AAH19345.1; -.
DR GO; GO:0005622; C:intracellular; IEA.
DR GO; GO:0004842; P:ubiquitin-protein ligase activity; IEA.
DR GO; GO:000512; P:ubiquitin cycle; IEA.
DR InterPro; IPR008973; C2:CaLB.
DR InterPro; IPR000569; HECT domain.
DR InterPro; IPR002349; WW_
DR InterPro; IPR001202; WW_rsp5_WWP.
DR Pfam; PF00632; HECT; 1.
DR Pfam; PF00397; WW; 4.
DR PRINTS; PR00403; WWDOMAIN.
DR SMART; SM00119; HECTC; 1.
DR SMART; SM00456; WW; 4.
DR PROSITE; PS00237; HECT; 1.
DR PROSITE; PS01159; WW DOMAIN 1; 4.
DR PROSITE; PS0020; WW DOMAIN 2; 4.
```

```
KW Hypothetical protein.
FT NON TER 1
SQ SEQUENCE 855 AA; 98604 MW; 3CF4A66996F033EA CRC64;

Query Match 31.6%; Score 1228.5; DB 4; Length 855;
Best Local Similarity 40.4%; Pred. No. 8.7e-22; Indels 93; Gaps 19;
Matches 276; Conservative 103; Mismatches 22;

Qy 93 LSNAGRL-----KDTGYQRLDLCKLNPSD-----TDVARGQ-----IVVSLQTRDR 134
Db 207 LSELSRLQITPDSNGEQFSSLIQREPPSRLRSCSVTDAVAEQGHLPSPSVAVVHTTPG 266
Qy 135 IGTGGSV-VDCRGLL-----ENEGTVESGPRPLSCFWEPPAPYTDSTGAAGGNN-- 186
Db 267 LPSGWEERKDAKGRYTVVNNHRTTWT-----RPIQLAEDGA-----SGSATNSNNHL 316
Qy 187 -----CRVEGPSQDORLQORLNPDPVGRSL-----OTPNRPHGHQSP----- 227
Db 317 IEQIRPRSLSTFTVTLSPLESAKDSPVRAVKDTLSNPQSPQSPYNSPKPKQKVQ 376
Qy 228 --LPEGVEQRTTVGGYVFLHTQGVSTWHDPRIP-----RDLNSVNCDELGPDPQWEV 280
Db 377 SFLPPGWEMEIAPNGRFFIDHNTKTTTWEDPRKLPFVHMRSKTSLNPNDLGPLPGWEE 436
Qy 281 RSTVSGRIYFVDHNNRTTQFTDPRLUHHIMHQCQLKEPSQPLPLPSEGSLEDEELPAQY 340
Db 437 RIHLDGRTFIDHNSKITQWEDPRLQN-----PAITG-----PAVPY 473
Qy 341 ERDLVQKLVRLHLSLQOPQAGHCRIEVSREEIFEEYSYROIIMK-RPKDLKKRLMYKPR 399
Db 474 SREFKQYDYFRKKLKKPADIPNRFEKMLHNNIFBESYRIRMSVKRPDVLKARLWIEFE 533
Qy 400 GEEGLDYGVAWEVLYLLCHEMLNPYYGLFOYSTNIYMLQINPDSSI-NPDHLSYPHFV 458
Db 534 SEKGLDYGVAWEVFLLSKEMFNYYGLFEYSATNYTLQINPNSGLCNEDHLSYPTFI 593
Qy 459 GRIMGLAVFHGYINGFTVPFYKQLLGKPIQLSDLESVDPELHKSILVILENDITPVL 518
Db 594 GRVAGLAVFHGKLDGFFIRFPYKMLGKQITLNDMESVDSEYVNSLKWILENDPTE-LD 652
Qy 519 HTFCVEHNAFGRILQHELKPNGRNVPVTEENKKEYVRLYVNRFRMGIEAQFLAQKFN 578
Db 653 LMFCIDSENFQTVQVDLKPNGSEIMVTENKKEYIDLVIQWRFVNRVQKMAFLEGFT 712
Qy 579 ELIPCHLLKPPDOKELELIIGLDKIDLNDWKSNTLRKHCVDNSN-IVRWFQAVETFE 637
Db 713 ELLPIDLIKFDENELELLMCGLDVDVNDWRQHSIYKNGYCNHPVIOQWFKAVLLMDA 772
Qy 638 ERRARLQFVTSRVPVLOGFKALQGSTGAAGRLFTIHLIDANTDNLKPAHCTCFNRDI 697
Db 773 EKIRLLQFVTGTSRVPVMPNGFAELYGSN---GPQLFTIEQW-GSPEKLPRAHTCFNRDL 828
Qy 698 PPYESYEKYLEKLITAVEETCGF 720
Db 829 PPYETFDLREKLLMAVENAQGF 851

RESULT 15
Q9BW58
ID Q9BW58 PRELIMINARY; PRT; 858 AA.
AC Q9BW58;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]_TaxID=9606;
RP SEQUENCE FROM N.A.
RC TISSUE=Skin;
RA Strausberg R.;
```


RL Submitted (NOV-2000) to the EMBL/GenBank/DBSJ databases.

DR EMBL; BC000621; AAH00621.1; -.

DR HSSP; Q13526; IPIN.

DR GO; GO:0005622; C:intracellular; IEA.

DR GO; GO:0004842; F:ubiquitin-protein ligase activity; IEA.

DR GO; GO:0006512; P:ubiquitin cycle; IEA.

DR InterPro; IPR008973; C2 CalB.

DR InterPro; IPR000569; HECT_domain.

DR InterPro; IPR002349; WW.

DR InterPro; IPR001202; WW_Rep5_WWP.

DR Pfam; PF00632; HECT; 1.

DR Pfam; PF00397; WW; 4.

DR PRINTS; PRO0403; WWDOMAIN.

DR SMART; SM00456; WW; 4.

DR PROSITE; PS0237; HECT; 1.

DR PROSITE; PS01159; WW DOMAIN_1; 4.

DR PROSITE; PS00020; WW DOMAIN_2; 4.

KW Hypothetical protein.

FT NON TER 1

SQ SEQUENCE 858 AA; 98918 MW; E2E3833F51162011 CRC64;

Query Match 31.6%; Score 1228.5; DB 4; Length 858;

Best Local Similarity 40.4%; Pred.No. 8.8e-92;

Matches 276; Conservative 103; Mismatches 211; Indels 93; Gaps 19;

QY 93 LSNAISRL-----KDTGYQLDLCKLNPSD-----TDAVRGQ-----IVVSLQTRDR 134

DB 210 LSELSRLQITPDSNGEQSSLIQREPSRLRSCSVTDAVQGHLPSPSVAYVHTTPG 269

QY 135 IGTGGSV-VDCRGLL-----ENECTVVEDSGPRLSCFMEEPAPYTDSTGAAGGCGN-- 186

DB 270 LPSCWEERKAKGRYYVYNNHNRITTTW----RPIMQLAEDGA-----SGSATNSNNHL 319

QY 187 -----CRFVESPSQORLQAORLNRDPVGRSL-----QTPQNRPHGHSPE----- 227

DB 320 IEPQIRPRSLSTVTLSAPLEGAKSPVRAVKDTLSNPQSPQSPYNSPKQHKVTQ 379

QY 228 --LPEGYEQRITVQGVYFHTQTGVSTWHDPRIP-----RDLNSVNCDELGLPPGWEV 280

DB 380 SFLPPGWMERIAPNGRPFFIDHNTKTTTWEDPRLKFPVHMRKSTLSLPNDLGLPLPGWEE 439

QY 281 RSTVSGRIYFVDHNNRTQFTDRLHLMHMQCOLKEPSQPLPLPSSEGSLEDEELPAQRY 340

DB 440 RIHLDGRTFYIDHNSKITQWEDPRLQN-----PATG-----PAVPY 476

QY 341 ERDLVQKLKVLRLHLSLQPOAGHCRIEVSREEIFEESYRQIMKM-RPKDLKKRLMYKFR 399

DB 477 SREFPKQYDYFRKKLKPADIPNRFEMKLRHNNIFEESYRIRMSVKRPPVLKARLWIEFE 536

QY 400 GEGLDYGGVAREWLYLICHEMLNPYYGLFOYSTDNTYMLQINPDSSI-NPDHLSYHFV 458

DB 537 SEKGLDYGGVAREWFFLLSKEMFNPYYGLFEYSATDNTYTLQINPNSGLCNEDHLSYFTFI 596

QY 459 GRIMGLAVFHGHYINGGTFVPFYKOLLGKPIQLSDLESVDPELHKSLSWLILENDITPVLD 518

DB 597 GRVAGLAVFHGKLDGDFIRFFYQWMLGKQITLNDMESVDSEYNSLKWILENDPTE-LD 655

QY 519 HTFCVENHAPGRIILQHELKPNRNVVPYVENKKEYRLYVNNRPMRGIEAQFLAOKGFN 578

DB 656 LMFCDIEENFGQTYQVDLKPNGSEIMVTNENKREYIDLVIQWRVNRVQKQNAFLLEGFT 715

QY 579 ELIPQHLKPPDOKELELILIGLGLDKILDNDKSNTRLKHCVADSN-IVRWFWQAVETFE 637

DB 716 ELLPIDLIKIFDELELELMCGLDGVDVNDNRQHSIYKNGYCFNHPVQWFKAVLLMDA 775

QY 638 ERRARLLQFVTGSTVPLQGFKALQSGTGAAGPRLFTIHLIDANTDNLPAKHTCFNRIDI 697

DB 776 EKIRLLQFVTGTSKVPNGFAELYGSN---GPQLFTIEQW-GSPEKLPRAHTCFNRDL 831

QY 698 PYESYEKLYEKLITAYEETCGF 720

DB 832 PPYETFDLREKLLMAVENAQGF 854

Search completed: September 21, 2004, 07:47:22
Job time : 96.9789 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 21, 2004, 07:47:37 ; Search time 91.9109 Seconds
(without alignments)
2526.134 Million cell updates/sec

Title: US-10-009-945-2

Perfect score: 3884

Sequence: 1 GSSIKIRLTVLCAKNLAKK.....EXLYKLLTAVETCGFAVE 723

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1342398 seqs, 321133274 residues

Total number of hits satisfying chosen parameters: 1342398

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA.*

- 1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep.*
- 10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
- 12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
- 13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
- 14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
- 15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
- 16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
- 17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
- 18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3878	99.8	722	14	US-10-097-534-14
2	3001	77.3	748	14	US-10-021-660-81
3	2996	77.1	804	12	US-10-220-120-407
4	2913.5	75.0	735	14	US-10-313-985-2
5	2236	57.6	514	11	US-09-764-875-819
6	1471	37.9	766	14	US-10-313-955-4
7	1436.5	37.0	832	14	US-10-032-585-7296
8	1345	34.6	869	14	US-10-128-714-8162
9	1259	32.4	927	14	US-10-097-534-15
10	1249.5	32.2	911	14	US-10-205-823-279
11	1245.5	32.1	739	12	US-10-183-936A-89
12	1245.5	32.1	739	14	US-10-097-534-10
13	1245.5	32.1	739	15	US-10-374-979-89
14	1245.5	32.1	732	15	US-09-919-039-235
15	1228.5	31.6	834	14	US-10-313-955-6

16	1228.5	31.6	995	14	US-10-097-534-9	Sequence 9, Appl
17	1228.5	31.6	995	14	US-10-205-823-275	Sequence 275, App
18	1227	31.6	725	14	US-10-185-050-126	Sequence 126, App
19	1226.5	31.6	759	14	US-10-128-714-3162	Sequence 3162, Ap
20	1226	31.6	854	14	US-10-205-823-277	Sequence 277, App
21	1225	31.5	870	14	US-10-097-534-12	Sequence 12, Appl
22	1225	31.5	906	14	US-10-185-050-48	Sequence 48, Appl
23	1223	31.5	854	14	US-10-287-218-3	Sequence 3, Appl
24	1223	31.5	854	16	US-10-474-291-3	Sequence 3, Appl
25	1212	31.2	474	10	US-09-774-639-371	Sequence 371, App
26	1212	31.2	474	10	US-09-369-730-249	Sequence 249, App
27	1212	31.2	474	16	US-10-621-363-249	Sequence 249, App
28	1195	30.8	898	12	US-10-188-186-114	Sequence 114, App
29	1047	27.0	683	14	US-10-185-050-46	Sequence 46, Appl
30	1047	27.0	684	14	US-10-097-534-11	Sequence 11, Appl
31	1023	26.3	380	14	US-10-307-956-32	Sequence 32, Appl
32	1012	26.1	375	14	US-10-307-956-31	Sequence 31, Appl
33	1002	25.8	1562	14	US-10-097-534-13	Sequence 13, Appl
34	836.5	21.5	1094	14	US-10-043-487-300	Sequence 300, App
35	836.5	21.5	1488	14	US-10-043-487-285	Sequence 285, App
36	836.5	21.5	2011	16	US-10-408-765A-767	Sequence 767, App
37	821.5	21.2	733	14	US-10-097-534-16	Sequence 16, Appl
38	807.5	20.8	1104	12	US-10-424-599-145002	Sequence 145002,
39	800	20.6	1843	12	US-10-424-599-145004	Sequence 145004,
40	776.5	20.0	3647	16	US-10-437-963-119793	Sequence 119793,
41	754.5	19.4	973	16	US-10-437-963-178922	Sequence 178922,
42	724	18.6	277	9	US-09-925-300-1628	Sequence 1628, Ap
43	720.5	18.6	358	14	US-10-268-036-5	Sequence 5, Appl
44	720.5	18.6	358	16	US-10-391-364-33	Sequence 33, Appl
45	691	17.8	157	12	US-10-211-462-107	Sequence 107, App

ALIGNMENTS

RESULT 1

US-10-097-534-14
; Sequence 14, Application US/10097534
; Publication No. US20030049607A1
; GENERAL INFORMATION:
; APPLICANT: GREENER, TSVIKA
; APPLICANT: MOSKOWITZ, HAIM
; APPLICANT: REISS, YUVAL
; APPLICANT: ALROY, IRIS
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE MODULATION OF VIRAL
; TITLE OF INVENTION: NATURATION
; FILE REFERENCE: PLY-001.01
; CURRENT APPLICATION NUMBER: US/10/097,534
; CURRENT FILING DATE: 2002-03-12
; PRIOR APPLICATION NUMBER: 60/275,224
; PRIOR FILING DATE: 2001-03-12
; PRIOR APPLICATION NUMBER: 60/308,958
; PRIOR FILING DATE: 2001-07-31
; PRIOR APPLICATION NUMBER: 60/340,170
; PRIOR FILING DATE: 2001-12-07
; NUMBER OF SEQ ID NOS: 71
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 14
; LENGTH: 722
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-097-534-14

Query Match 99.8%; Score 3878; DB 14; Length 722;

Best Local Similarity 100.0%; Pred. No. 0; Mismatches 0; Indels 0; Gaps 0;

Matches 722; Conservative 0;

Qy 2 GSSIKIRLTVLCAKNLAKKFFRLPDPFAKIVDGSQCHSTDTVKNLDPKMNQHYDLY 61

Db 1 GSSIKIRLTVLCAKNLAKKFFRLPDPFAKIVDGSQCHSTDTVKNLDPKMNQHYDLY 60

Qy 62 VKGTDSTISVNNKHTKKQAGFLGCVRLLSNAISRLKDTGYORLDLCKLNESDIDAV 121

```
Db 61 VGKTDSTIISVNNHKKIHKQAGFLGCVRLLSNAISRLKDTGYQRLDCLKNPSTDVA 120
Qy 122 RGOIVSLOTDRIGTGSVVDRCGLLENGTUVYDSCGPRPLSCFWEPPAPYTDSTGAA 181
Db 121 RGOIVSLOTDRIGTGSVVDRCGLLENGTUVYDSCGPRPLSCFWEPPAPYTDSTGAA 180
Qy 182 AGGNCRFVESPSQORLOAQRLENPDVRGSLQTPQNRPHGHSQPELPEGYEQRTTVQGG 241
Db 181 AGGNCRFVESPSQORLOAQRLENPDVRGSLQTPQNRPHGHSQPELPEGYEQRTTVQGG 240
Qy 242 VYFLHTQGVSTWHDPRIPRLNSVNCDELCPPLPGHVEVSTSGRIYFVDHNNRTTQFT 301
Db 241 VYFLHTQGVSTWHDPRIPRLNSVNCDELCPPLPGHVEVSTSGRIYFVDHNNRTTQFT 300
Qy 302 DPLHHMHQOLKEPESQPLPLPSEGSLDEELPAQRYERDLVQKLKVLRLHLSLQOPQ 361
Db 301 DPLHHMHQOLKEPESQPLPLPSEGSLDEELPAQRYERDLVQKLKVLRLHLSLQOPQ 360
Qy 362 AGHCRIVSRBEIPEESYRQIMKWRPKDLKRLMKVFRGBEGLDYGAVREWLKLCHEM 421
Db 361 AGHCRIVSRBEIPEESYRQIMKWRPKDLKRLMKVFRGBEGLDYGAVREWLKLCHEM 420
Qy 422 LNPYVGLFQYSDNLYMLQINPDSSINPDHLSYFHFVGRIMGLAVFHHYINGGFTVPFY 481
Db 421 LNPYVGLFQYSDNLYMLQINPDSSINPDHLSYFHFVGRIMGLAVFHHYINGGFTVPFY 480
Qy 482 KOLGKPIQLSDLESVDPELHKSLLWILENDITPVLDTHTFCVEHNAFGRILQHELKPNR 541
Db 481 KOLGKPIQLSDLESVDPELHKSLLWILENDITPVLDTHTFCVEHNAFGRILQHELKPNR 540
Qy 542 NVPVTEENKKEVRLVNNRFRVGI EAOFLALOKGENELIPQHLKPPDOKELELIIGGL 601
Db 541 NVPVTEENKKEVRLVNNRFRVGI EAOFLALOKGENELIPQHLKPPDOKELELIIGGL 600
Qy 602 DKIDLNDWKSNTLKHCVADSNIVRWFQAVETDEERRARLLQFVTGSTRVPLQGFAL 661
Db 601 DKIDLNDWKSNTLKHCVADSNIVRWFQAVETDEERRARLLQFVTGSTRVPLQGFAL 660
Qy 662 QGSTGAAGPRFLTTHLIDANTDNLPAKHTCFNRDIDIPPYESYKLYEKLTAVEETCGFA 721
Db 661 QGSTGAAGPRFLTTHLIDANTDNLPAKHTCFNRDIDIPPYESYKLYEKLTAVEETCGFA 720
Qy 722 VE 723
Db 721 VE 722

RESULT 2
US-10-021-660-81
; Sequence 81, Application US/10021660
; Publication No. US20030152926A1
; GENERAL INFORMATION:
; APPLICANT: Murray, Richard
; APPLICANT: Glynn, Richard
; APPLICANT: Watson, Susan R.
; TITLE OF INVENTION: EOS Biotechnology, Inc.
; FILE OF INVENTION: Compositions and Methods of Screening for Angiogenesis,
; TITLE OF INVENTION: Modulators
; FILE REFERENCE: 018501-0007100S
; CURRENT APPLICATION NUMBER: US/10/021,660
; CURRENT FILING DATE: 2001-12-06
; PRIOR APPLICATION NUMBER: US/09/784,356
; PRIOR FILING DATE: 2001-02-14
; PRIOR APPLICATION NUMBER: US 09/637,977
; PRIOR FILING DATE: 2000-08-11
; NUMBER OF SEQ ID NOS: 135
; SOFTWARE: FastSeq for Windows Version 3.0.
; SEQ ID NO 81
; LENGTH: 748
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-021-660-81
```

```
Query Match 77.3%; Score 3001; DB 14; Length 748;
Best Local Similarity 74.6%; Pred. No. 5,5e-266;
Matches 565; Conservative 63; Mismatches 71; Indels 58; Gaps 8;

Qy 5 IKIRLTVICAKMLAKKDFRLLPDPFAKIVVDGSGCHSTDTVNTLDPKWNQHDVLYGK 64
Db 12 VKLRLTVICAKMLAKKDFRLLPDPFAKIVVDGSGCHSTDTVNTLDPKWNQHDVLYGK 71
Qy 65 TDSITISVNNHKKIHKQAGFLGCVRLLSNAISRLKDTGYQRLDCLKNPSTDVAVRGQ 124
Db 72 SDSVITISVNNHKKIHKQAGFLGCVRLLSNAISRLKDTGYQRLDCLKNPSTDVAVRGQ 131
Qy 125 IVVSLQTRDRIGTGSVVDRCGLLENE-----GTYY----- 155
Db 132 IVVSLQTRDRIGTGSVVDRCGLLENE-----GTYY----- 191
Qy 156 --EDSGPRPLSCFWEPPAPYTDSTGAAAGGNCRFVESPSQORLOAQRLENPDVRGSL 213
Db 192 ASGEYSPGRPLSCFWEPPAPYTDSTGAAAGGNCRFVESPSQORLOAQRLENPDVRGSL 242
Qy 214 QTFQNRPHGHSQPELPEGYEQRTTVQGVYFLHTQGVSTWHDPRIPRLNSVNCDELGP 273
Db 243 ---SRTHLHTPPLPEGYEQRTTVQGVYFLHTQGVSTWHDPRIPRLNSVNCDELGP 298
Qy 274 LPPGHEVSTVSGRIYFVDHNNRTTQFTDPR-----LHHMHQOLKEPESQPLPLPSEG 329
Db 299 LPPGHEVSTVSGRIYFVDHNNRTTQFTDPR-----LHHMHQOLKEPESQPLPLPSEG 354
Qy 330 L---EDELPAQRYERDLVQKLKVLRLHLSLQOPQAGHCRIVSRBEIPEESYRQIMKWR 386
Db 355 LCPDDETECLTVPRYKEDLVQKLKVLRLHLSLQOPQAGHCRIVSRBEIPEESYRQIMKWR 414
Qy 387 PKDLKRLMKVFRGBEGLDYGAVREWLKLCHEMNLNYPYGLFQYSTNLYMLQINPDSS 446
Db 415 PKDLKRLMKVFRGBEGLDYGAVREWLKLCHEMNLNYPYGLFQYSTNLYMLQINPDSS 474
Qy 447 INPDHLSYFHFVGRIMGLAVFHHYINGGFTVPFYKQLGKPIQLSDLESVDPELHKSLLV 506
Db 475 VNPVTEENKKEVRLVNNRFRVGI EAOFLALOKGENELIPQHLKPPDOKELELIIGGL 534
Qy 507 WILENDITPVLDTHTFCVEHNAFGRILQHELKPNRNVPTTEENKKEVRLVNNRFRVGI 566
Db 535 WILENDITPVLDTHTFCVEHNAFGRILQHELKPNRNVPTTEENKKEVRLVNNRFRVGI 594
Qy 567 EAOFLALOKGENELIPQHLKPPDOKELELIIGGLDKIDLNDWKSNTLKHCVADSNIVR 626
Db 595 EAOFLALOKGENELIPQHLKPPDOKELELIIGGLDKIDLNDWKSNTLKHCVADSNIVR 654
Qy 627 WFWQAVETDEERRARLLQFVTGSTRVPLQGFALQGSTGAAGPRFLTTHLIDANTDNL 686
Db 655 WFWQAVETDEERRARLLQFVTGSTRVPLQGFALQGSTGAAGPRFLTTHLIDANTDNL 711
Qy 687 KAHTCFNRDIDIPPYESYKLYEKLTAVEETCGFAVE 723
Db 712 KAHTCFNRDIDIPPYESYKLYEKLTAVEETCGFAVE 748

RESULT 3
US-10-220-120-407
; Sequence 407, Application US/10220120
; Publication No. US20040048253A1
; GENERAL INFORMATION:
; APPLICANT: INCYTE GENOMICS, INC.
; APPLICANT: PANZER, Scott R.
; APPLICANT: SPIRO, Peter A.
; APPLICANT: BANVILLE, Steven C.
; APPLICANT: SHAH, Purvi
; APPLICANT: CHALUP, Michael S.
; APPLICANT: CHANG, Simon C.
; APPLICANT: CHEN, Alice
; APPLICANT: D'SA, Steven A.
; APPLICANT: AMSHEY, Stefan
```



```

TELEPHONE: (617) 832-1000
TELEFAX: (617) 832-7000
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
    LENGTH: 735 amino acids
    TYPE: amino acid
    TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO:
US-10-313-955-2

```

[illegible]

RESULT 5
US-09-764-875-819
; Sequence 819, Application US/09764875
; Publication No. US20040018969A1
; GENERAL INFORMATION:

```

; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: P202
; CURRENT APPLICATION NUMBER: US/09/764,875
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 1249
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 819
; LENGTH: 514
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-764-875-819

```

Query Match	57.6%	Score 2236	DB 11	Length 514
Best Local Similarity	80.7%	Pred. No. 5.6e-196		
Matches 414	Conservative 43	Mismatches 42	Indels 14	Gaps 4
QY	218	NRPHCHQSPPELPEGEYORTTVQGGVYFLHTQTGVSHTWHDPR:PRDLNSVNCDELGFLPPG	277	
DB	9	SRTHLHTPPDLPEGEYORTTQGGVYFLHTQTGVSHTWHDPRVPRDLNSINCELGFLPPG	68	
QY	278	MEVRTVSGRYFYVDHNNRRTTQTDPR:---LHHIMNHQCOLKERSQPLPSEGL---	330	
DB	69	HEIRNTATGRVYFYVDHNNRRTTQTDPRLSANLHLVLRNQQLKQQQQQV:---SLCPD	124	
QY	331	EDEELPAQRYERDVLQKLKVLRHLSLQOPAGCHRIEVSREEIFEEVSRYQIMKWRPKDL	390	
DB	125	DTECLTVPRYKRDVLQKLKLRBELSQOPAGCHRIEVSREEIFEEVSRYQIMKWRPKDL	184	
QY	391	KKRLMKVRGEGGLDYGAVAREWMLYLLCHEMLNPPYGLFOYSTDNILYQLINPDSINPD	450	
DB	185	WKRLMIKRGEGGLDYGAVAREWMLYLLSHEMLNPPYGLFOYSRDRIYTLQINPDSAVNPE	244	
QY	451	HLSYFHFVGRIMGVAFHGHVINGGFTVPYKQLLGKPIQLSDLSVDPQLHKSILVWILE	510	
DB	245	HLSYFHFVGRIMGVAFHGHYIDGGFTLPYKQLLGKSIITLDMELVDPDLHSLVWILE	304	
QY	511	NDITPVLDTFCVEHNAFGRILQHELKPNGRNVPTTEKNKKEYVLYVNWPMRGIEAQF	570	
DB	305	NDITGVLDHTFCVEHNAYGEBI:QHELKPNKGS:IPVNEENKKEYVLYVNWPMRLRGIEAQF	364	
QY	571	LALQGENELTPOHLKPFQCKELEJTIIGLKDKIDANKSITLKHCVADSNIVWPMQ	630	
DB	365	LALQGFNEVITPOHLKLTDFEKELEJTIIGLGKIDVNDKVNTRLKHCTPDSNIVKWEWK	424	
QY	631	AVEFDFEERRARLLQFVTGSTRVPLQGFKAQGSTGAAGPRLFTTHLIDANTDNLPKAHT	690	
DB	425	AVEFDFEERRARLLQFVTGSSRVPLOGFKAQ:---GAAGPRLFTTHQIDACTNNLPKAHT	481	
QY	691	CNFRIDIPPPYESKYLEKLLTAVEETCGFAVE	723	
DB	482	CNFRIDIPPPYESKYLEKLLTAIEETCGFAVE	514	

```

RESULT 6
US-10-313-955-4
; Sequence 4, Application US/10313955
; Publication NO. US20030199036A1
; GENERAL INFORMATION:
; APPLICANT: Beach, David H.
; Caligiuri, Maureen
; Nefsky, Bradley
; TITLE OF INVENTION: Ubiquitin Ligases, and Uses Related Thereto
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FOLEY, HOAG & ELIOT LLP
; STREET: One Post Office Square
; City: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109-2170
; COMPUTER READABLE FORM:

```


	Prior Application Number:	60/340,170	
	Prior Filing Date:	2001-12-07	
	Number of Seq ID NOS:	71	
	Software:	PatentIn Ver. 2.1	
	Seq ID NO 15	:	
	Length:	927	
	Type:	PRT	
	Organism:	Homo sapiens	
	US-10-097-534-15		
	Query Match	32.4%; Score 1259; DB 14; Length 927;	
	Best Local Similarity	33.6%; Pred. No. 8,1e-106;	
	Matches	307; Conservative 129; Mismatches 239; Indels 240; Gaps 28;	
QY	7 IRLTVLCANVLAKKDPFLPDPFAIV---VDGSGQHSTDTVKNLDPKW---	55	
DB	48 VVRVIAGIGLAKKOILGASDYVTVDPMNGVLTSVQTTKTSLNPKWNEELFRV	107	
QY	56 ---QH-----YDLVGXKD--SITISVN-----HKIKHKGOG 83		
DB	108 HQOHLLFEVDENLRTRDDFLGVQDVPLPTPENRLEPRPYTFKFVHLHRSHRSRV	167	
QY	84 AFLGCGRULLSNARLKXT-----GYORLD-----LCKL-----	113	
DB	168 KGYL-----RLKMITYLPKTSGSDNAEQAELEGFWLVLPDPAACHLOQQOE	216	
QY	114 -----NPED--TDVARGQIWSIQ-----TRDR 134		
DB	217 PSELPPCWEERQDIIGRYYNHESRRTCWKRPFPDNLTDAENGNIQLQAQAFTRRQ	276	
QY	135 IGFGSVVPDCRGLEN-----EGTYVEDGG-FGRPLSCFMEEPPAPYT-	176	
DB	277 ISBETSDVNQSSENWEIRREDEATMYSSQAFTSPPSNLNDVPTHABELNARLTIFG	336	
QY	177 ----STGAAGGNCR-----IVESPS-----QQRLOAQL-- 204		
DB	337 NSAVOPASSNHSGRRGLAQYTEEOFTLVLLTTSSLGPLPGWEKEQGERGSYYVDH	396	
QY	205 -----RNPDVGRSLQPQ-----NRPHGHOSPELPEGYEQ 234		
DB	397 NSRTTWTPKPVAIVEITSOLTSOSSAGPOSASTSDSQOVTOPSEIEQGFLPKGVEV	456	
QY	235 RTVVQCQVFHTGTGVSWHDPR--IPRDLSNV----NCDELGPPLPGWEVRTVSRI	288	
DB	457 RHAPNGRPFDIHNTKITWTWDPRLIKPAHLRGKTLSDLTNSDLGLFPLPGWEERTHDGI	516	
QY	289 YFVDHNRRNTQTDDLRLHHMHOCLEKPSOPLPLPSEGSLDEDLPACRYERDLVKKL	348	
DB	517 FYINNIKOTQEDPLENV-----AITGPAV-----YSIDYKRKY 553		
QY	349 KVLRHLSLQPOAGCHRIEVSREBIFFESTYQIMMKMPKD-LKKRLMVYPGEGLDYG	407	
DB	554 EPPERKLKQNIDPNKFKBRATVLEDSTRYMINGVRADFLKRWLFDFDGKGLDYG	613	
QY	408 GVAREWLILLCHEMLNPHYGLFOYSTDNITVMLOINPSSI-NPDHLSYEHFVGRIMGVAY	466	
DB	614 GVAREWFILSKEMFNPIYGLFEYSATNYTLQINPNGLCNEDHLSYFXFYGRVGMAY	673	
QY	467 FHGYINGGFYPFYKOLLGKIQLDSLEVDPHELHKSGLVMILENDITPVLDHTFCVEHN	526	
DB	674 YHGKLLDGGFIHFYKOWMLKHKIIITHDMESVDSEYNSLRMLENDPTE-LDLRFIDE	732	
QY	527 AGRILIQHELNKGRNVPTYENKEYVLYNWRFMKGIAQFLAQKGFNELIPOHLL	586	
DB	733 LFQOTHCLKMGGSSEIVVTNNKKEYTYLIQWRFNVRIOQMAAFKEGFFEIPQOLI	792	
QY	587 KPDPQELELIIGGLDKIDLNDWNKSNTRLKH-CVADSINI VWFWOAVENTDEERRALLQ	645	
DB	793 KIFDENELLEMCGLGDVDVNDWRHTHYKVGSYSAHQVIOFWKAVLMMDSEKELLQ	852	
QY	646 FVTGSTRVLPQGFALQCGTAGGRLPTTIHLIDANTONLPCAHCNFNRIDIIPPYSYEK	705	
DB	853 FVTGSRVPMWGFAELYGSN---GPSSTVEGW-GTPEKLPRACTCFNRILDLPPTSFE	908	
QY	706 LYEKLLTAVBETECPG 720		
DB	909 LWDKLOMAIENTQGF 923		
	RESULT 10		
	US-10-205-823-279		
	; Sequence 279, Application US/10205823		
	; Publication No. US20030108963A1		
	; GENERAL INFORMATION:		
	; APPLICANT: Schlegel, Robert		
	; APPLICANT: Monahan, John E.		
	; APPLICANT: Endege, Wilson O.		
	; APPLICANT: Ganavarapu, Manjula		
	; APPLICANT: Gorbatcheva, Bella		
	; APPLICANT: Hoersch, Sebastian		
	; APPLICANT: Kamatkar, Shubhangi		
	; APPLICANT: Wonsay, Angela M.		
	; APPLICANT: Glatt, Karen		
	; APPLICANT: Zhao, Xumei		
	; APPLICANT: Anderson, Dustin		
	; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND		
	; METHOD FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND		
	; THERAPY OF PROSTATE CANCER		
	; FILE REFERENCE: MEI-044		
	; CURRENT APPLICATION NUMBER: US/10/205,823		
	; PRIOR FILING DATE: 2002-07-25		
	; PRIOR APPLICATION NUMBER: 60/307,982		
	; PRIOR FILING DATE: 2001-07-25		
	; PRIOR APPLICATION NUMBER: 60/314,356		
	; PRIOR FILING DATE: 2001-08-22		
	; PRIOR APPLICATION NUMBER: 60/325,020		
	; PRIOR FILING DATE: 2001-09-25		
	; PRIOR APPLICATION NUMBER: 60/341,746		
	; PRIOR FILING DATE: 2001-12-12		
	; PRIOR APPLICATION NUMBER: 60/362,158		
	; PRIOR FILING DATE: 2002-03-05		
	; NUMBER OF SEQ ID NOS: 455		
	; SOFTWARE: FastSEQ for Windows Version 4.0		
	; SEQ ID NO 279		
	; LENGTH: 911		
	; TYPE: PRT		
	; ORGANISM: Homo sapiens		
	US-10-205-823-279		
	Query Match	32.2%; Score 1249.5; DB 14; Length 911;	
	Best Local Similarity	34.5%; Pred. No. 5.9e-105;	
	Matches	303; Conservative 120; Mismatches 243; Indels 211; Gaps 24;	
QY	17 LAKXDFFRLP---DPPAKI---VDGSGQCH---STDVTNKTLDPKWOHYDLYVGTKD-S 67		
DB	69 LPISDELFLPGICDPYVKLSLYVADENRELALVQTKTKIKTNPKWNEEFYRVNPSNHR	128	
QY	68 ITISVNWKKHXKQAGFLGCVRLSNSAI-----SERLQDT 103		
DB	129 LLFVDEFNLTRDD---FLGQDVDPLSHLPDETDMERPYPYTFKOFLLRPRSCHKSRVK--	183	
QY	104 QYORLDLCLKNPSDTDAVRCQIVVSLQTRDIGTGGSVVDCRG-----	146	
DB	184 GFRLKWAYMPCNG-----GQDEENSQDQDDMEHGWEVVDNSDSASQHOELEPPPPLP	238	
QY	147 ---LLENEGTVY-----	157	
DB	239 WEEKVDNLGRTRYVYVHNMTTOM		


```

; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-097-534-10

Query Match      32.1%; Score 1245.5; DB 14; Length 739;
Best Local Similarity 39.8%; Pred. No. 9.8e-105;
Matches 274; Conservative 99; Mismatches 180; Indels 135; Gaps 18;

QY      114 NPQSDTAVRQIVVSL-----QTRDRIG-TGGSVVDCRGLLENBGTVEYDSGGPRLSCF 167
      :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      109 SPATSSDGSSTGSLPPTNTNTSSGATSGLIPL-----TISGGSGP-RPLNPFV 159
      :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY      168 MEHPAPYTDGTGAAAGGNCRFVESPQDQRLOAQRLRNPDVRGSL-----CTPQNR 219
      :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      160 TQAPLP-----PQWQRV-----DQGRVYVDHVEKRTTWR 192
      :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY      220 PHGQSPPEGVQRVTVGQVFLHTQTGVSTWHDPRIP----- 260
      :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      193 PE-----PLPFGWERRVDNNGRIYVDFHFRITTTWQRTLESVRNVEQWLQRSLOGAM 244
      :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY      261 -----RDJ-----NSVNCDELGLPPGMEVRSTVSGRIYFVDHNRRTQFTDPR 305
      :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      248 QQNRFIYQNDLFATSQSKFDPGLPGLPPGMEKRTDSGRVYFVNHNTRITQWEDPR- 306
      :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY      306 HHMNHQCLKEPSQPLPL-----PSEGSLEDEELPAQR 339
      :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      307 -----SQQLNE--KPLPEGWEMRFTVDGIPYFVDHNRRTTYIDPRTGKSALDNGPOIA 359
      :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY      340 YERDLVQKLVLKVR---HELSQLQPAQACHCRLEVSBEIPEESYRQIMKMRPKDLKKELMV 396
      :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      360 YVRDFKAKVQYFRFCQQLAMPQ---HIKLTVTRKTLFEDSPQOIMSFSPQDLURRLDW 415
      :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY      397 KFRGEGLDGGVAREMYLLCHEMLNPYVYGLFOYSTDNIMYLQINPDSSINPDHLSYFH 456
      :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      416 IFFGEGLDGGVAREWFFLLSHEVLNPMYCLPEYAGKDNVCLQINPASYINPDHLYFR 475
      :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY      457 FVGRIMGLAVPHGYINGGTVPYKOLLGKPIOLSDLESVDPELHKSLLVILENDITPV 516
      :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      476 FIGFTAMALFHGKFIDTGFSLPFPYKRLNKPVLGKXLESIDPEFYNLSLWYKNNTEEC 535
      :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY      517 -LDHTFCVEHNAFGRIILCHELKPNGRNVPTVEENKKEYVRLVYNVNRMRGIEAFLALQK 575
      :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      536 DLEWYFSDVEILGETIKSHDLKPNGGNILLVTEENKEYIRNVAEWRLSRGVEEQTAFFE 595
      :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY      576 GFNELIPQHLKPPDQKLELIITGLDKIDLNDWKSNTRLKHGCVADSNIVRWFMQAVET 635
      :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      596 GFNEILPQYLOVFDKALEVLLCGMCEIDLNDWQRAIYRHVARTSKQIMWFWQFVKEI 655
      :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY      636 DEERRALLOFVGTSTVPVLCGFKALOGSGAAGPRLFTIHLIDANTDNLPKAHTCNRI 695
      :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      656 DNEKRMELLOFVTGCTPLPVGFGFADLMGSN---GPQFCIEKV-GKENWLPRLSRHTCFNRL 711
      :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY      696 DIPPEYSYKLYKLLTAVETTCGFAYE 723
      :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      712 DLPPKSYEQKKEKLLFAIETEGFQGE 739
      :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

RESULT 13
US-10-374-979-89
; Sequence 89, Application US/10374979
; Publication No. US2003021979A1
; GENERAL INFORMATION:
; APPLICANT: John P. Carulli et al.
; TITLE OF INVENTION: THE HIGH BONE MASS GENE OF 11q13.3
; FILE REFERENCE: 032796-021
; CURRENT APPLICATION NUMBER: US/10/374,979
; CURRENT FILING DATE: 2003-03-04
; PRIOR APPLICATION NUMBER: US 09/544,398
; PRIOR FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: US 09/543,771
; PRIOR FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: US 09/229,319
; PRIOR FILING DATE: 1999-01-13

```


Db	573	GRVAGLAVFHGKLLDGGFFIRPFYKMLGKQITLNDMESVDSEYNSLKWILENDPTE-LD	631
QY	519	HTFCVEHNAFGRILQHELKPNGRNVPVTENKKEYVRLVYNNRFRMGIEAQFLALOKGN	578
Db	632	LMFCIDEENFGQTYQVDLKNPNSSEIWTNENKREYIDLVIQWRFNVRVQKQWNAFLEGET	691
QY	579	ELIPQHLKPFDOKELELIIGGLKIDLNDWKSNTRLKHCVADSN-IVRWFQAVETFE	637
Db	692	ELLPIDLIKIFDENELELLMCGLDVDVNDWRQHSYKNGYCNHPVIOFWKAVLLMDA	751
QY	638	EREARLLQVTSRVPLOGFKALQSTGAAGPRLETHILIDANTDNLKHAHTCFNRDI	697
Db	752	EKIRLLQFVTGSRVPMNGFAELYSN---GPQLFTISQW-GSPEKLPRAHTCFNRDL	807
QY	698	PPYESYKLYEKLLTAVEETCGF	720
Db	808	PPYETEDLREKLLMAVENAQGF	830

Search completed: September 21, 2004, 08:04:23
Job time : 99.4109 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 21, 2004, 07:39:16 ; Search time 30.0014 Seconds
(without alignments)
2398.267 Million cell updates/sec

Title: US-10-009-945-4
Perfect score: 4038
Sequence: 1 MSNPGRRNGPVKLELTWLC.....EKLYEKLTAETTCGPAVE 748

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_78.*
1: pir1.*
2: pir2.*
3: pir3.*
4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1585.5	39.3	767	2	T37545
2	1564	38.7	766	1	ubiquitin-protein
3	1521	37.7	815	2	probable ubiquitin
4	1484	36.8	809	1	ubiquitin-protein
5	1418	35.1	786	2	ubiquitin-protein
6	1364.5	33.8	820	2	ubiquitin-protein
7	1355.5	33.6	887	2	ubiquitin-protein
8	1177	29.1	708	2	NBDD-4 ORF - mouse
9	1046.5	25.9	671	2	ubiquitin ligase N
10	858	21.2	3227	2	probable ubiquitin
11	846.5	21.0	3839	2	probable ubiquitin
12	805.5	19.9	1126	2	related to TOM1 pr
13	803	19.9	4056	2	ubiquitin-protein
14	789	19.5	3268	2	protein F14J16.10
15	709.5	17.6	889	2	hypothetical prote
16	626	15.5	310	2	DNA binding protei
17	596	14.8	221	2	hypothetical prote
18	534.5	13.2	874	2	Be-associated prot
19	465	11.5	807	2	probable ubiquitin
20	460.5	11.4	892	2	probable membrane
21	454	11.2	873	2	polyubiquitin-like
22	449	11.1	1142	2	hypothetical prote
23	445.5	11.0	1054	2	hypothetical prote
24	444.5	11.0	1083	2	hypothetical prote
25	418.5	10.4	1029	2	probable ubiquitin
26	407.5	10.1	4861	2	giant protein p619
27	403.5	10.0	1066	2	hypothetical prote
28	387	9.6	1001	2	hypothetical prote
29	371	9.2	4836	2	herc2 protein - mo

30 360 8.9 655 2 T41750
31 358 8.9 757 2 T05688
32 355.5 8.8 1502 2 T48309
33 342 8.5 910 2 S64155
34 297.5 7.4 2899 2 T21546
35 297.5 7.4 2915 2 G87867
36 286.5 7.1 959 2 T14761
37 281 7.0 889 2 S22659
38 278.5 6.9 1483 2 S30015
39 262.5 6.5 632 2 T38617
40 245.5 6.1 2761 2 T29285
41 224.5 5.6 2895 2 T08437
42 220.5 5.5 1630 2 T00390
43 207 5.1 472 2 B56954
44 204 5.1 1256 2 JE0209
45 197 4.9 1171 2 T42372

ALIGNMENTS

RESULT 1

T37545
ubiquitin-protein ligase (EC 6.3.2.19), pub1 [validated] - fission yeast (Schizosaccharomyces pombe)
N;Alternate names: ubiquitin ligase Pub1
C;Species: Schizosaccharomyces pombe
C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 03-Jun-2002
C;Accession: T37545; T48655
R;McLean, J.; Harris, D.; Barrell, B.G.; Rajandream, M.A.; Wood, V.
submitted to the EMBL Data Library, September 1997
A;Reference number: Z21722
A;Accession: T37545
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-767 <MCL>
A;Cross-references: EMBL:Z99161; PIDN:CAB16207.1; GSPDB:GN00066; SPDB:SPAC11G7.02
A;Experimental source: strain 972h-; cosmid c11G7
R;Saleki, R.; Jia, Z.; Karagiannis, J.; Young, P.G.
Mol. Gen. Genet. 254, 520-528, 1997
A;Title: Tolerance of low pH in Schizosaccharomyces pombe requires a functioning pub1 ubi
A;Reference number: Z07985; MUID:97340937; PMID:9197411
A;Accession: T48655
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-767 <SAL>
A;Cross-references: EMBL:U62795; NID:G2262192; PIDN:AAB63350.1; PID:G2262193
A;Experimental source: strain J227
C;Genetics:
A;Gene: pub1; SPDB:SPAC11G7.02
A;Map position: 1
A;Intons: 6/2; 14/1; 62/2
C;Function: <CYC>
A;Description: EC 6.3.2.19 [validated, MUID:96205868]
C;Function: <CYC>
A;Description: Involved in of the mitotic activating tyrosine phosphatase cdc25 [validate
C;Function: <TOU>
A;Description: required for low pH-tolerance [validated, MUID:97340937]
C;Superfamily: yeast ubiquitin-protein ligase, WW repeat homology
C;Keywords: cell cycle control; ligase
F;205-242/Domain: WW repeat homology <WW1>
F;288-325/Domain: WW repeat homology <WW2>
F;345-382/Domain: WW repeat homology <WW3>

Query Match 39.3%; Score 1585.5; DB 2; Length 767;

Best Local Similarity 41.3%; Pred. No. 9.8e-107;
Matches 340; Conservative 121; Mismatches 231; Indels 131; Gaps 17;

QY 1 MSNPGRRNGPVKLELTWLCARLVKDFRLLPDPFAKVVVDGSGQCHSTDTVKNTLDPK 60

Db 1 MSNSAQSR----RIKRVITVAADGLYKRDVFRFPDPFAVLTVDGE-QTHTTIAKKTLPY 55

QY 61 VNHQYDLYIGKSDSVTISVWVNHKKHKQAGFLGCVLFLSNRLNRLKDTGYQL--DLG 118

Db 56 WNETFEVNTDNTSTIAIQVDFQKKE-KKGGQGLGVINLRVGVLDLAIGDEMTRDLK 114
 QY 119 KLGPNNDTVRGQIVVSQ----- 137
 Db 115 K--SNENTVVGKIIINLSTTAOSTLQVPSAASGARTQRTSITNDPQSSQSSVSERNPA 172
 QY 138 -----SRDRIGTGGQVDCSRLFDN-----DLPDGWEERTASGRIOYLNIHITRTQ 184
 Db 173 SSRAGSPTRDPAAPASSEPFTSSFDQYGLRPFGEWERTDNLGRTYVVDHNTSTT 232
 QY 185 WERTP-----RPASEYSSGRPLSCFVDENTPISGTNGATCGOSSDPLAERRVRSORH 238
 Db 233 WIRPNLSSVAGAAAEHLSSAS--SANVTGQVQSSNAA-----RTEASVLT--- 279
 QY 239 RNYMSRTHLTPDLPPEGYEORTTCQGVYFLHTQTVSTWHDPRVPRDLSNIN----- 292
 Db 280 ----SNATTAGSGELPPGWEQRYTPGGRPYFVDHNTTRITTTWDPRRQYIRSYGGPNNA 335
 QY 293 -----CEELGPLPPGWEIRNTATGRVYFVDHNNRTTQDPRLSANLHLVLRQNLKQD 347
 Db 336 IQQPVSQLGPLPSGWEMLNTARVYFVDHNTKTITWDDPRPSSL-----DQ 384
 QY 348 QQQVVSCLPDDTECLTVPRYKRDLVOKLILRQELSQ--QPQAGHCRIEVSREIPEE 405
 Db 385 -----NVQYKRDPRKLIYF---LSQPALHPLPGQCHIKVRNHIPE 425
 QY 406 SYRQVVKORPKDLWKRLMTKFRGEGDYGGAWEMLYLLSHEMLNPPYGLFOYSRDDIY 465
 Db 426 SYAEIMRQSATDLKRLMTKFDGEDGLDYGLSREYFLLSHEMFPNPFYCLPEYSSVDNY 485
 QY 466 TLQINPDSAVNPEHLSVFEHVGRIQMGMAVFGHYIDGGFTLFFYKOLLGKSLTLDMLV 525
 Db 486 TLQINPHSGINPEHLYYFVIGRVLGIAIFHRFVDVAFVWSFYKMLQKKTLDQMESM 545
 QY 526 DPDLHNSLVWLENDITGVLDHTFCVHNAYGEIIOHELKPNKSGIPVNEENKKEYVRLY 585
 Db 546 DAEYRSLVWLDNDITGVLDITFTSVEDNCFGEVVTIDLPKNGRIEVTENKREYVDLV 605
 QY 586 VNRFLRGIEAQFALQKGFNEVPIQHLKLTDEKELELIICGLGKIDVNDKVNRLKH 645
 Db 606 TVW-IQRIEEQNAPHEGFSELIPELINVDFERELELLIGGISEIDMEDWKKHVDYS 664
 QY 646 CTPODSNTVKWFWKAVFEDEERRARLQVFGSSRVPLQGFALQGAAGPRFLTTHQIDA 705
 Db 666 YSENDQIKFWFELMDWESNEKSRLLQTTGTSIPVNGFXDQSGDGRKFTIEKAGE 725
 QY 706 CTNNLPKAHTCFNRIDIPPYESYKLYEKLTAIBETCGFAVE 748
 Db 726 -PNKLPKAHTCFNRLLDPPYTSKDLHKLSTAVETIGFGQE 766
 RESULT 2
 S66562
 N:ubiquitin-protein ligase (EC 6.3.2.19) - fission yeast (Schizosaccharomyces pombe)
 N:Alternate names: E6-AP-like protein ubiquitin ligase
 C:Species: Schizosaccharomyces pombe
 C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 03-Jun-2002
 C:Accession: S66562; T45159
 R:Nefsky, B.; Beach, D.
 EMBO J. 15, 1301-1312, 1996
 A:Title: Publ acts as an E6-AP-like protein ubiquitin ligase in the degradation of cdc25
 A:Reference number: S66562; MUID:96205868; PMID:8635463
 A:Accession: S66562
 A>Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-766 <NEP>
 A:Cross-references: GB:U66716; NID:g1519443; PIDN:AAB07514.1; PID:g1519444
 R:Nefsky, B.S.; Beach, D.
 submitted to the EMBL Data Library, August 1996
 A:Description: Publ acts as an E6-AP-like protein ubiquitin ligase in the degradation of
 A:Reference number: Z22935
 A:Accession: T45159
 A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA
 A:Residues: 1-766 <NE2>
 A:Cross-references: EMBL:Y07592; PIDN:CAA68867.1
 C:Genetics:
 A:Gene: publ
 C:Superfamily: yeast ubiquitin-protein ligase; WW repeat homology
 C:Keywords: ligase
 F:205-242/Domain: WW repeat homology <WW1>
 F:288-325/Domain: WW repeat homology <WW2>
 F:345-382/Domain: WW repeat homology <WW3>
 Query Match 38.7%; Score 1564; DB 1; Length 766;
 Best Local Similarity 41.1%; Pred. No. 3 Se-105;
 Matches 338; Conservative 121; Mismatches 232; Indels 132; Gaps 18;
 QY 1 MSNPRRNRGPKLRLTVLCAKVLKDKDFRLLPDPAPKVVGGSGCHSTDTVKTLDPK 60
 Db 1 MSNSAQSR---RIRVTIIVAADGLYKRDVFRFPDPFAVLVDGE-QTHTTTAKKTLNPY 55
 QY 61 WNOHYDLYIGKSDSVTISVWNNKKIHKQKAGFGLGCVRLLSNAINRLKDTGYORL--DLC 118
 Db 56 WNETFEVNTDNTSTIAIQVDFQKKE-KKGGQGLGVINLRVGVLDLAIGDEMTRDLK 114
 QY 119 KLGPNNDTVRGQIVVSQ----- 137
 Db 115 K--SNENTVVGKIIINLSTTAOSTLQVPSAASGARTQRTSITNDPQSSQSSVSERNPA 172
 QY 138 -----SRDRIGTGGQVDCSRLFDN-----DLPDGWEERTASGRIOYLNIHITRTQ 184
 Db 173 SSRAGSPTRDPAAPASSEPFTSSFDQYGLRPFGEWERTDNLGRTYVVDHNTSTT 232
 QY 185 WERTP-----RPASEYSSGRPLSCFVDENTPISGTNGATCGOSSDPLAERRVRSORH 238
 Db 233 WIRPNLSSVAGAAAEHLSSAS--SANVTGQVQSSNAA-----RTEASVLT--- 279
 QY 239 RNYMSRTHLTPDLPPEGYEORTTCQGVYFLHTQTVSTWHDPRVPRDLSNIN----- 292
 Db 280 ----SNATTAGSGELPPGWEQRYTPGGRPYFVDHNTTRITTTWDPRRQYIRSYGGPNNA 335
 QY 293 -----CEELGPLPPGWEIRNTATGRVYFVDHNNRTTQDPRLSANLHLVLRQNLKQD 347
 Db 336 IQQPVSQLGPLPSGWEMLNTARVYFVDHNTKTITWDDPRPSSL-----DQ 384
 QY 348 QQQVVSCLPDDTECLTVPRYKRDLVOKLILRQELSQ--QPQAGHCRIEVSREIPEE 405
 Db 385 -----NVQYKRDPRKLIYF---LSQPALHPLPGQCHIKVRNHIPE 425
 QY 406 SYRQVVKORPKDLWKRLMTKFRGEGDYGGAWEMLYLLSHEMLNPPYGLFOYSRDDIY 465
 Db 426 SYAEIMRQSATDLKRLMTKFDGEDGLDYGLSREYFLLSHEMFPNPFYCLPEYSSVDNY 485
 QY 466 TLQINPDSAVNPEHLSVFEHVGRIQMGMAVFGHYIDGGFTLFFYKOLLGKSLTLDMLV 525
 Db 486 TLQINPHSGINPEHLYYFVIGRVLGIAIFHRFVDVAFVWSFYKMLQKKTLDQMESM 545
 QY 526 DPDLHNSLVWLENDITGVLDHTFCVHNAYGEIIOHELKPNKSGIPVNEENKKEYVRLY 585
 Db 546 DAEYRSLVWLDNDITGVLDITFTSVEDNCFGEVVTIDLPKNGRIEVTENKREYVDLV 605
 QY 586 VNRFLRGIEAQFALQKGFNEVPIQHLKLTDEKELELIICGLGKIDVNDKVNRLKH 645
 Db 606 TVW-IQRIEEQNAPHEGFSELIPELINVDFERELELLIGGISEIDMEDWKKHVDYS 664
 QY 646 CTPODSNTVKWFWKAVFEDEERRARLQVFGSSRVPLQGFALQGAAGPRFLTTHQIDA 705
 Db 666 YSENDQIKFWFELMDWESNEKSRLLQTTGTSIPVNGFXDQSGDGRKFTIEKAGE 724
 QY 706 CTNNLPKAHTCFNRIDIPPYESYKLYEKLTAIBETCGFAVE 748
 Db 726 -PNKLPKAHTCFNRLLDPPYTSKDLHKLSTAVETIGFGQE 766

RESULT 3

T49744
probable ubiquitin-protein ligase [imported] - Neurospora crassa
N:Alternate names: protein B24B19.160
C:Species: Neurospora crassa
C>Date: 02-Jun-2000 #sequence_revision 02-Jun-2000 #text_change 18-Aug-2000
R:Accession: T49744
R:Schults, U.; Aign, V.; Hohseil, J.; Brandt, P.; Fartmann, B.; Holland, R.; Nyakatura,
submitted to the Protein Sequence Database, May 2000
A:Reference number: Z5022
A:Status: preliminary
A:Accession: T49744
A:Molecule type: DNA
A:Residues: 1-815 <SCH>
A:Cross-references: EMBL:AL356182; GSPDB:GN00116; NCSP:B24B19.160
A:Experimental source: BAC clone B24B19; strain OR74A
C:Genetics:
A:Gene: NCSP:B24B19.160
A:Map position: 6
A:Introns: 11/1, 24/1, 59/2, 110/1, 783/2
C:Superfamily: yeast ubiquitin-protein ligase; WW repeat homology
F:239-276/Domain: WW repeat homology <WWR2>
F:334-371/Domain: WW repeat homology <WWR2>
F:393-430/Domain: WW repeat homology <WWR3>
Query Match 37.7%; Score 1521; DB 2; Length 815;
Best Local Similarity 38.8%; Pred. No. 5, 1e-102; Indels 170; Gaps 20;
Matches 332; Conservative 124; Mismatches 229;
QY 16 LTVLCANLVKDPFRLPDPFPAKVVDGSGCHSTDTVKNTLDPKNOHYDLY-----I 69
DB 9 LPVIAADGLYKRDVPRFPDPAVATNGE-QTKTQVSKRTLPVYVNEHDFWLIQSKV 67
QY 70 GKSDSVTT:SVNHHKHKHKKOGAGFLGCVRL-LSNAINELKDTGYCRL--DLCKLGPNDND 126
DB 68 NEDSLAVQVDFQKAF-KKKQGGFLGVINRGVDELAPDAEDQMLTRDLKK--STDNL 124
QY 127 TVRGQIVVSLQ-----SRDRI-----GTG 145
DB 125 VVHGKLIINLTNTATWSRLGPPSSSRPSLLTPQSSVINDRANERPSSAMSGPNTA 184
QY 146 GOVVDCSR-----LPDND--LPDG 162
DB 185 NNMTLASSPASLAVSSSSSTAPTGTNGTAPNTPTSLVPAQARHSTLSLPEDSMGLPAG 244
QY 163 WEERTASGRQYLNHITRTTQWERPTPASEYSRPSRPLSCFVDENTPI----- 212
DB 245 WERREDHLGRYYVDHNSRTTSWNEPT-----GTGAENRTAEANTQVERQHQRNRL 297
QY 213 ----SGTNGATCGSSDPRLAERRVRSORHRYNYSRTHLHTP--PDLPGEYEQRTTQQGQ 266
DB 298 PEDRTGAN-----SPTLQQQAAATANAAATMMHTGATTPGTGELPAGWEQRTPEGR 349
QY 267 VFLLHTQGVSTWHDPRVPRDL-----SNIN-----CDELGLPDPGHEIRNTATGRV 313
DB 350 PYFDVHNTRTTWTWDPRQYIRMYGGONNTNGTIQQQPVSQLGLPLSGMELTNTARV 409
QY 314 YFVDHNRRTTQDTPRLSANLHLVLRNQNQLKQOQQQVWSLCPDDTECLTVPRYKRDV 373
DB 410 YFVDHNTKTTTWDPRLPSSL-----DQ-----NVQYKEDR 442
QY 374 QKLKILRQLSQOQPOACHCEIEVSREIPFESTRYQWKPDPDLWRLMIKFRGEGSLD 433
DB 443 RKLIFRSQ-PAMRIMSQCCHIKVRRSHIFEDSPAEISROSATDLKRLMIKFDGEDGLD 501
QY 434 YGVVAREWLYLLSHEMLNPPYGLFYGRDRIYTIQINPDGAVNPEHLSYFHFVGRIMGMA 493
DB 502 YGGLSREFFFLSHMEMNPFYCLEYSADHNTYTIQINPHSGINPEHLNYPKFTGRVVGLA 561
QY 494 VFHGHIYDGGTLPFYKQLKSKSTLDDMELVDPDLNLSVLWLENDITGVLDHTFCVEH 553
DB 562 IFRFLDPAFFIGALYKMWLGAVALDMEGVDADFRSLQWMLDNDITDVLDAFTFSD 621
QY 554 NAYGEIIQHEUKPKNGKSPVYVNEENKKEVRLVYVNRFLRGIEAQFLALQKGFNEVIPH 613

DB 622 ERFGVITEEDLIENGRIAVTNNKKEVYELVMVKWRIEKIEQGFRAFQGHFELIPQDL 681
QY 614 LKTFDEKELLLIICGLKIDVNDKWNTRLKHCTPPDSNIVKWFKAVEFDFDEERRALLQ 673
DB 682 INVDFESELELLIGGTAEDVDWKKHTDYRGYTESDEVDFWQTVRSWDGSKRLLQ 741
QY 674 FVTGSSRVPLQGFKAQGAAGPRFLFTIHQIDACTNNLPAKHTCFNRIDIPPPYSYEKLYE 733
DB 742 FTTGTSRIPVNGPKDLOGSDGPRFTIERKAGEIT-NLPKHTCFNRLDLPFYKSLMLQ 800
QY 734 KLITATBETCGPAVE 748
DB 801 KLITAVEETMGFGQE 815
RESULT 4
S43217
ubiquitin-protein ligase (EC 6.3.2.19) RSP5 - yeast (Saccharomyces cerevisiae)
N:Alternate names: E6-AP-like protein ubiquitin ligase; protein YER125w; PUB1 protein
C:Species: Saccharomyces cerevisiae
C>Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 03-Jun-2002
C:Accession: S43217; S50628; S70050
R:Mulligan, J.T.; Dietrich, F.S.; Hennessey, K.M.; Sehl, P.; Komp, C.; Wei, Y.; Taylor, I.
submitted to the EMBL Data Library, February 1993
A:Reference number: S30812
A:Accession: S43217
A:Molecule type: DNA
A:Residues: 1-809 <MUL>
A:Cross-references: GB:U18916; EMBL:L11119; NID:gl384128; PIDN:AAC03223.1; PID:g603364
R:Dietrich, F.S.
submitted to the EMBL Data Library, December 1994
A:Description: The sequence of S. cerevisiae cosmid 9781, 8198, 9115, 9981, and lambda c
A:Reference number: S50628
A:Accession: S50628
A:Molecule type: DNA
A:Residues: 1-809 <DIE>
A:Cross-references: EMBL:U18916; NID:gl384128; PIDN:AAC03223.1; PID:g503364; GSPDB:GN0000
R:Hein, C.; Springael, J.Y.; Volland, C.; Haguenauer-Tsapis, R.; Andre, B.
Mol. Microbiol. 18, 77-87, 1995
A:Title: NP11, an essential yeast gene involved in induced degradation of Gap1 and Fur4
A:Reference number: S70050; MUID:96154942; PMID:8596462
A:Accession: S70050
A:Status: nucleic acid sequence not shown; not compared with conceptual translation
A:Molecule type: DNA
A:Residues: 1-101 <HEI>
C:Genetics:
A:Gene: SGD:RSP5; PUB1; NP11; MIPS:YER125W
A:Cross-references: MIPS:YER125W; SGD:S0C00927
A:Map position: 5R
C:Function:
A:Description: involved in endocytosis of GAP1 protein and FUR4 protein; binds and ubiqui
C:Superfamily: yeast ubiquitin-protein ligase; WW repeat homology
C:Keywords: ligase
F:229-266/Domain: WW repeat homology <WW1>
F:331-368/Domain: WW repeat homology <WW2>
F:387-424/Domain: WW repeat homology <WW3>
Query Match 36.8%; Score 1484; DB 1; Length 809;
Best Local Similarity 37.1%; Pred. No. 2, 4e-99;
Matches 324; Conservative 125; Mismatches 224; Indels 200; Gaps 20;
QY 11 PVKRLRTVLCAKNLVKDPFRLPDPFPAKVVDGSGCHSTDTVKNTLDPKNOHYDLY-YI 69
DB 2 PSSISVKLVAAESLYKRDVFRSPDPFAVLITDGY-QTKSTSAKTLNPNYNETKFDI 60
QY 70 GKSDSVTISVWVHHKHKHKKOGAGFLGCVRL-LSNAINRLKQ-----TGVR-----LDLC 118
DB 61 NENSILTIQVFPQKXF-KKKDQOQFLGVNVRVGVLDHLEDATSSGRPRETITRDLK 119
QY 119 KLGPNDNDTVRGQIVV-----SLQSRDRIGTGQVVDCSR--- 153
DB 120 K-SNDGMVSGRLIVLVLSKLPSSSPHQAQPSGHTASSSTTSSTRTRNGHSTSTRNHS 177

154 QY -----LFDND-----LPDGNEE 165
 178 DB TGHPSRGTAQAVESTLQSCCTAATNTATTSHRSTNSTSATQYSSFFEDQVGRLPFGWER 237
 166 QY RTASGRICVLAHITTTTOWERPSPASE-----YSSPGRLPSCFVDENT 210
 238 DB RTDNFGRTYVDHNTTRITTKRPTLDTQTEAERGQNLNANTELERRQHRGRTLPGSSSDNS 297
 211 QY -----PISG-----TNGATCGOSSDPRLAERRVRSQRHNNYMSHTLHT 249
 298 DB SVTVQVGGGNSIPVNGAAAAAFAATGTTSLGL 331
 250 QY PPDLPGEYQORTQOQVFLHTQCVSTWHDPRVPRDLNIN-----CEELGFL 299
 332 DB --ELPGWQRFPPEGRAYFVDHNTTITWVPRQOYRTYGPNTTIQQQPVQLGFL 389
 300 QY PGWEIRNTATGRVYFVDHNNRTTQFTDPRLSANLHLVLRNQNLKQDQOQOQVSLCPDD 359
 390 DB PSCEWMLNTARVYFVDHNTKTTTWDPRLPSSL-----DQ----- 426
 360 QY TECLTVPRYKRDVLQKILRQELSQQOQ-----AGHCIEVSRBEIPEESYRQVWKMRP 415
 427 DB ---NVPOYKRDPRKVIYFR-----SOPALRILPQCCHIKVRKKNIFEDATQEMRQTP 477
 416 QY KDLWKRLMIKFRGEBGLDYGVARWMLYLLSHEMLNPPYGLFYQSRDDIYTLQINPDSAV 475
 478 DB EDLKKRLMIKFDGEBGLDYGVSREFFLLSHEMFNFFYCLFEYSAYDNYTIQINPNSGI 537
 476 QY NPEHLSYFHVGRIMGVPHGYHIDGGFTLFPYKOLLGKSLTLDLMELVDPDLHNSLWV 535
 538 DB NPEHLNFKYIGRWVGLGVFHRFLDAFFVGLAYKMLRKKVYLQDMEGVDAEYNSLWV 597
 536 QY ILENDITGLDHTFCVEHNAYGELIQHELKPNKSIPIVNEENKKEVYRLVYVNRFLRGIE 595
 598 DB MLENSIDGVLDTLFSADDERFGEVTVVDLKPDRNLEVDGNKKEVYELVYQWLRVQ 657
 596 QY AQFLALQKGFNEVPIQHLKTFDEKELELIICGLGKIDVNDKVNTRLKHTCPDSDNIVKW 655
 658 DB EQKAFMDGFNELIPEDLVTVFDERELLELLIGIAEIDIEDWKHTDYRGYQESDEVIQW 717
 656 QY FWKAVEFDEBRARLLOFTVTSRIPVNGFKDLQSGDGRPRRTIEKAGE-VQQLPKSHT 715
 718 DB FWKCVSEWDEQARLLOFTTTSRIPVNGFKDLQSGDGRPRRTIEKAGE-VQQLPKSHT 776
 716 QY CFNRIDIPPPYESYEKLEKLLTAIETCGFAVE 748
 777 DB CFNRVDTLPOYVDYSKQKLTLAVEETIGFGQE 809
 RESULT 5
 T39585
 ubiquitin protein ligase - fission yeast (Schizosaccharomyces pombe)
 C:Species: Schizosaccharomyces pombe
 C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 18-Aug-2000
 C:Accession: T39585
 R:Voickaert, G.; Wood, V.; Rajandream, M.A.; Barrell, B.G.
 submitted to the EMBL Data Library, August 1997
 A:Reference number: Z21865
 A:Accession: T39585
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-786 <VOL>
 A:Cross-references: EMBL:Z99759; PIDN:CAB16903.1; GSPDB:GN00067; SPDB:SPBC16E9.11c
 A:Experimental source: strain 972h-; cosmid c16E9
 C:Genetics:
 A:Gene: SPDB:SPBC16E9.11c
 A:Map position: 2
 A:Introns: 60/2
 C:Superfamily: yeast ubiquitin-protein ligase; WW repeat homology
 F:236-273/Domain: WW repeat homology <WWR1>
 F:306-343/Domain: WW repeat homology <WWR2>
 F:364-401/Domain: WW repeat homology <WWR3>

Query Match 35.1%; Score 1418; DB 2; Length 786;
 Best Local Similarity 36.7%; Pred. No. 1.4e-94;
 Matches 316; Conservative 127; Mismatches 217; Indels 202; Gaps 20;
 10 GPVKRLTVLCARNLVKKOPFRLPDPFAKVVVDGSCQCHSTDTVKNLDPKKNQHYDLVI 69
 4 GAKRVRYIIVAADGLSKRDLFRQDPFAITVDGE-QTHTKVIKSVNPNWNEGEFVTV 62
 70 GKSDSVTISVWVHKIKKAGAGFLGCVLLSNAINRLKDTGYQRDL-----CKLGP 122
 63 KPSVISIRLFDOKKP-KKKDQGLGLVSP-----RMREVGSFRNREVSRLTRPKKSS 115
 123 NDNDTVRGQIVV----- 134
 116 TTNLVLGNLKVAPSKIRAPAGNHSSTTANRTTSTPTTTTARTTRTPRPTATTNS 175
 135 -----SLQSRDRIGTGG-----QVVDCSRLFDND----- 158
 176 QSTNSNTRNGTSAATNGTGTGAGTGASHRSPPVTRQTNNTSALNSNAHINMSFEDQY 235
 159 --LPDGWEERTASGRIOYLNHITRTTOWERPTRPASEYSSPCRLSCFVDENTPISGNT 216
 236 GRLLPGWERRADSLGRYYVDHNTRTITW--TRPAS-----STNPVHNT- 277
 217 GATCGQSSDPLAERRVRSQRHNNYMSRTHLHTPP-----DLPEGEVQORTQ 263
 278 -----SSD-----SOR-LNHQNR---HLDDSNPMSQSDSGNDLFFGWEMRYTD 318
 264 QGVYFLHTTOTGVSTWHDPR-----VPRDLNINCEELGPLPPGWEIR 306
 319 TGRPYFVDHNTTRITTWVDPNPLVRPNNGSGSTVGLMQPOSLSH-----LGPLSPGWEMR 373
 307 NTATGRVYFVDHNNRTTQFTDPRLSANLHLVLRNQNLKQDQOQOQVSLCPDTECLTVP 366
 374 LTNSARVYFVDHNTKTTTWDPRLPSSAL-----DQD-----VP 406
 367 RYKEDLVQKILRQELSQQOQPAQGHCRIEVSRBEIPEESYRQVWKMRPKDLWRLMIKF 426
 407 QYKCDFRKLYFASQ-PGMRPLPGQCNVVRDHIPEDSYAEIMRYSAHDLKRLMIRF 465
 427 RGEELDYGVARWMLYLLSHEMLNPPYGLFYQSRDDIYTLQINPDSANVPEHLSYFHV 486
 466 DGEDGLDYGVSREFFLLSHKMPDIYCLFEYSADVNTYLIQINPHSSINPEHLNFRFI 525
 487 GRIMGMAVFHGYHIDGGFTLFPYKOLLGKSLTLDLMELVDPDLHNSLWLENDITGLD 546
 526 GRVIGLAIFHRFLDAFFVSVLYKLLRKKVSLADMESIDAEPYRSLKWLWLENDITGLD 585
 547 HTFCVHNAYGEITQHCLKPNKSIPIVNEENKKEVYRLVYVNRFLRGIEAQFLALQKGFN 606
 586 LTFSEEDHFGVEVTEVLTNGENIEVTEENKKYVDLVTEWVSRKRVQOQNAFVSGFV 645
 607 EVIPOHLKTFDEKELELIICGLGKIDVNDKVNTRLKHTCPDSDNIVKWEKAVEFFDEE 666
 646 ELVSPDLVNVFDERELLELLIGISVDVEDWKSHTYRTVIATDPVVIKFWELIAGWNE 705
 667 RRARLLOFTVTSRIPVNGFKALQAGAPRLFTTHQIDACTNNLPKARTCFNRIDIPPEY 726
 706 DRKLLQFATGTSRIPVNGFADLQSGDGRPKFTIEKA-GTPDQLPVAHTCFNRLLDLPYP 764
 727 SYEKLVEKLLTAIETCGFAVE 748
 765 SKDTLHEKLSLAVENTVGVFGNE 786
 RESULT 6
 T46412
 ubiquitin-protein ligase (EC 6.3.2.19) NEDD4 - human (fragment)
 N:Alternate names: hypothetical protein DKF2p434p2422.1
 C:Species: Homo sapiens (man)
 C:Date: 04-Feb-2000 #sequence_revision 04-Feb-2000 #text_change 03-Jun-2002
 C:Accession: T46412

R:Blum, H.; Bauersachs, S.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.
submitted to the Protein Sequence Database, January 2000
A:Reference number: Z23034
A:Accession: T45412
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-820 <AAA>
A:Cross-references: EMBL:AL137469
A:Experimental source: adult testis; clone DKFZp434P2422
C:Genetics:
A:Gene: GDB:NEDD4
A:Note: DKFZp434P2422.1
C:Superfamily: rat ubiquitin-protein ligase; protein kinase C C2 region homology; ubiquitin
C:Keywords: ligase
F:142-179/Domain: WW repeat homology <WWR1>
F:342-379/Domain: WW repeat homology <WWR2>
F:393-430/Domain: WW repeat homology <WWR3>
F:489-814/Domain: ubiquitin-protein ligase homology <UBI>

Query Match 33.8%; Score 1364.5; DB 2; Length 820;
Best Local Similarity 37.1%; Pred. No. 1.2e-90;
Matches 315; Conservative 115; Mismatches 221; Indels 199; Gaps 25;
QY 50 TDTVNTLDPKQNHVDYLYIGKSD-SVTISVWNHKKHKKQAGFLGCVR----- 98
Db 12 TKTIKTLNPKWNEEFYFRVNPNSHRLLFVFDENLRTDD---FLGQVDVPLSHLPTD 68
QY 99 -----LLSNAINLKDTCYQRLDLCKLGNPDNDTVRGQIVVSLOSRRDRTGIG 146
Db 69 PTMERPYTFKDFLLPRSHKSEVKFLKXAYMPNG-----GDENSRQDDMEHW 123
QY 147 QVDCSRLLFDND-----LPDGWEHRTASGRIOYLNHITTTQWERPT----- 189
Db 124 EYVD-----SNDASQHOBELPPPLPPGWEKVDNLGRTYVNHNNRTQWHRPSLMDV 178
QY 190 -----RPASEYSSP-----GRPLSCFVDENTPISGTN-- 216
Db 179 SSESNNTRQINQEAHRRFRSRHISDLPEPSEGGDVEPNET-ISEENVNAGDSLG 237
QY 217 -----GATCGOSSDPR-LAE-----RRVR-----SORHRYNMRTHL----- 247
Db 238 LALPPPPASPGSRTSQBELSELSRLQ:TPDSNGEQSSLIQREPSKSLSCSVTDVA 297
QY 248 ---HTPPD-----LPGEYEQRTQCGQ 266
Db 298 EGHLLPPGAKSPVRRAVYKTLNPNQSPQSPYNSPKQHKVQTSFLPPGWMRIAPNGR 357
QY 267 VYFLHTQGVSTWHDPRVP-----RDLNINCEELGPPGWEIRNTATGRVYFDHNR 321
Db 358 PFFIDHNKTTTWEDPRLKFPVHESKTSLNPNDLGPPGWEIRHLDCRTFFIDHNSK 417
QY 322 TQFTDPLSLANHLVLNRQNLKDDQQQQVVSCLPDDTECLTVPR--YKDLVQKLIL 379
Db 418 ITQWEDPRL-----QN-----PAITGPAVPSREFKQKYDYF 449
QY 380 RQLSQQQPQAGCHRIEVSREBIPEESYQVMKVRPKDLWK-RLMKIPRGEGLDYGVA 438
Db 450 RKLLKPADIPNRFEMKLRHNNIFESYRRINMSVKRPDVLKARLWIEFSEKGDYGV 509
QY 439 REWLYLLSHEMLNPPYGLFQYSRDDIYTLQINPSAV-NPEHLSYFHPFVGRIMGAVFH 497
Db 510 REWFLLSKEMFNPPYGLFEYSATDNYTLQINPSGLCNEHLSYFTFGRVAGLAVFH 569
QY 498 HYIDGGFTLPVKQLGKSLTLDDELVDPLHNSLWILENDITGLVDHTFCVEHNAYG 557
Db 570 KLLDGGFTTRPFXKMLGKQITLNDNESVDSEYNSLRKILENDPT-ELDLMFCDENFG 628
QY 558 EIIQHELKPKNGKSIPIVNEENKKEYVRLVYNNRFLRGIEAFLQCKGFNEVPOHLLKTF 617
Db 629 QTYQVDLKPNGSEIMVTNENKREYIDVIOKRFVNRVCKQNNAFLEGFTTELLPDLIKIP 688
QY 618 DEKELELLICGLKIDVNDKVNTRLK--HCTPDSNVKFWKVAVEFDEERRARLLQFV 675

Db 689 DENELELLMCGLDGVDVNDHRQHSIYKNGYC-PNHFVIOFWKAVLMDAEKRIRLLQFV 747
QY 676 TGSRRVPLQGFKALQGAAGPRLFTIHIQIDACTNNLPAKHTCFNRIDIPPIYESYKLEKL 735
Db 748 TGTSRVPMNGFAELYGSGNPGOLFTEIQWGS-PEKLPRHTCFNRLDLPYTFEDLREKL 806
QY 736 LTAIEETCGF 745
Db 807 LMAVENAQQF 816
RESULT 7
S70642
ubiquitin ligase Nedd4 - rat (fragment)
C:Species: Rattus norvegicus (Norway rat)
C>Date: 14-Feb-1997 #sequence_revision 13-Mar-1997 #text_change 18-Aug-2000
C:Accession: S70642
R:Staub, O.; Dho, S.; Henry, P.C.; Correa, J.; Ishikawa, T.; McGlade, J.; Rotin, D.
EMBO J. 15, 2371-2380, 1996
A:Title: WW domains of Nedd4 bind to the proline-rich PY motifs in the epithelial Na(+) c
A:Reference number: S70642; MUID:96221297; PMID:8665844
A:Accession: S70642
A:Status: preliminary; nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 1-887 <STA>
A:Cross-references: EMBL:U50842; NID:gl293646; PIDN:AAB48949.1; PID:gl293647
C:Genetics:
A:Gene: Nedd4
C:Superfamily: rat ubiquitin-protein ligase; protein kinase C C2 region homology; ubiquitin
F:54-167/Domain: protein kinase C C2 region homology <KC2>
F:246-283/Domain: WW repeat homology <WW1>
F:402-439/Domain: WW repeat homology <WW2>
F:459-496/Domain: WW repeat homology <WW3>
F:555-881/Domain: ubiquitin-protein ligase homology <UBI>
Query Match 33.8%; Score 1355.5; DB 2; Length 887;
Best Local Similarity 37.1%; Pred. No. 5.8e-90;
Matches 314; Conservative 131; Mismatches 248; Indels 153; Gaps 25;
QY 14 LRLTVLCAKMLVKDFRLLPDPPEAKVY---VDGSOCHSTDVTKVTLDPKMNCHVDLY- 68
Db 77 VRVKYIAGLAKKDLGASDPYRVTLDPMSGLTVSVQTKIKSLAPKNEEI-LFR 135
QY 69 -IGKSQSVTISVWNHKKHKKQAGFLGCVR-----LLSNAIN 105
Db 136 VLPQQRHILFEVDENLRTDD---FLGQVDVPLPPTENPRMERPYTFKDFVLHPRSH 192
QY 106 RLKDTGVQRLDLCKLGNPDNDTVRGQIVVSLOSRRDRTGIGQVVD---CSRLL----FDN 157
Db 193 KSRVKGVLRLKMTYLPNGSDD-----ENADQAELEPGWVVDQPDAAATHLQHPPEPS 246
QY 158 DLPDGWEHRTASGRIOYLNHITTTQWERPT-----RDLNINCEELGPPGWEIRNTATGRVYFDHNR 193
Db 247 PLPPGWEERQDLGRTYVYVNHESRTTQWRKPSPEDDLTDENGDIQLQAHGATTTTRQIS 306
QY 194 E-----YSSPQRPPLSCFVDENTPISGTNATCGOSSDP-----RLAE----- 230
Db 307 EDVGDGNHESPENWEIVREDENTYISG-QAVQSPSPSGHPDVQVRLAESLDTLRLTWYGNP 365
QY 231 ---RRVRSORH--RNYMSRT-----HLHTPPDLPGEVQORTTQGGVYFLHTQT 274
Db 366 ATSQPTSSNHSRSGSSQTCIFEEQPTLPVLPTSSGLPPGWEEKQDGRSYYVDHNS 425
QY 275 GVSTWHDPP-----RVP-----RDLNINCEELGPPGWEIRNTATGRVYFDHNRRT 323
Db 426 KTTTWSKPTMQDDPRSKIPAHLRGKTFPDSNDLGPLPPGWEERTHTDGRVFFNNHKK 485
QY 324 QFTDPLSLANHLVLNRQNLKDDQQQQVVSCLPDDTECLTVPRYKRDVLVQKLILRQL 383
Db 486 QWEDPRM-----QNVAITGPAEP-----YSRDYKRYEYFRRKL 519
QY 384 SQQQPQAGHCRIEVSREBIPEESYQVMKVRPKDLWK-ELNFKRGEGLDYGVAWEVL 442

Db 520 KKQTDIPNPKPMLRRANILEDSSYRIMGKRAADFLKARLWIEFDGEGKLDYGGVAREWF 579
 QY 443 YLISHMLNPPYGLFOYS-RDDYITLIQINPDSAV-NPEHLSYFHFVGRIMGMAVTHGHYI 500
 Db 580 FLISKEMFNPPYGLFEYSATNYITLIQINPNSGLCNEDHLSYFKFGRVAGNAVHVKLL 639
 QY 501 DGGFTLPFFYKQLGKSGITLDDMELVDPDLHNSLVWILENDITGVLDHTFCVHNAYGEII 560
 Db 640 DGFIRPFYKQMLQKLITLHDMESVDSEYSSLRWILENDPT-ELDLRFIIDEELFGQTH 698
 QY 561 QHLEKPKNGKSIPIVNEENKKEYVLYVNWRFVLRGIEAQLALOKGENEVIPOHLLKTFDEK 620
 Db 699 QHLEKTKGSEIVVTNKNKKEYIYLVQWRFNRIQKMAAFKEGFFELIPQDLIKIFDEN 758
 QY 621 EELIIICGLGKIDVNDKVNTRLKH-CTPDSNIVKFWKAVFFDERRARLLQFVTGSS 679
 Db 759 EELIIMCGLGDVDNDWREHTKYNGYSNMHQVHFWKAVLWMDSEKRIQLLQFVTGTS 818
 QY 680 RVPLOGFKALQAGAPRLFTIHOIDACTNNLPKATCFNRIDIPPYESYEKLYEKLITAI 739
 Db 819 RVPMMNGFAELYGSGPQSFVEQW-GTPDKLPRACHTCFNRDLPPYESFDELWDKLQWAI 877
 QY 740 EETCGF 745
 Db 878 ENTQGF 883

RESULT 8
 NEDD-4 ORF - mouse (fragment)
 C:Species: Mus musculus (house mouse)
 C>Date: 02-Aug-1996 #sequence_revision 02-Aug-1996 #text_change 18-Aug-2000
 C/Accession: 183196
 R:Kumar, S.; Tomooka, Y.; Noda, M.
 Biochem. Biophys. Res. Commun. 185, 1155-1161, 1992
 A:Title: Identification of a set of genes with developmentally down-regulated expression
 A:Reference number: 160167; MUID:92328780; PMID:1378265
 A/Accession: 183196
 A/Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-708 <RES>
 A/Cross-references: GB:D10714; NID:g220508; PID:g220509
 C:Genetics:
 A:Gene: NEDD-4
 C:Superfamily: rat ubiquitin-protein ligase; protein kinase C C2 region homology; ubiquitin
 F:40-77/Domain: WW repeat homology <WW1>
 F:196-233/Domain: WW repeat homology <WW2>
 F:251-288/Domain: WW repeat homology <WW3>
 F:347-682/Domain: ubiquitin-protein ligase homology <UBI>

Query Match 29.1%; Score 1177; DB 2; Length 708;
 Best Local Similarity 40.4%; Pred. No. 3.6e-77;
 Matches 260; Conservative 87; Mismatches 178; Indels 118; Gaps 17;

QY 159 LPDQWEERTASGIQVNIHTRTQWERPT-----RPASE 194
 Db 42 LPQWEERQDVLGRYTVVNHESRTTKWRPSDDLDLTDNDMDQLQAQAFTRRQISE 101
 QY 195 -----YSSGRPLSCFVDENTPISG-----TNGATCGQ--S 223
 Db 102 DVDGPDNRESNWEIVRENTSEYQAVQSPSGHIDVQTHLASEFNRLAVCGNPAT 161
 QY 224 SDPRLAERVRVSQRHRYMGRTH-----LHTPPDLPEGYEQRTTQGGQYIFL 270
 Db 162 SQP-----VTSSNHSRGGSLQTCIFEEQPTLPVLLPTSSGLPPGWEERQDDRGSRYYV 215
 QY 271 HTQGVSTWDDP-----RVPKDL-SNINCEELGRLPGWEIRNTATGRVYFVDHNR 321
 Db 216 DHNSKTTTWSKPTWDDPRSKIPAHLRGKTDNDLGLPGLPWEERTHTDGRVFFIHNK 275
 QY 322 TTQFTDPRLSANLHLVLRNQLKDDQQQQVSLCPDDECTELTPVRYKRLDVQKLILRQ 381
 Db 276 KTQWEDPRL-----QNVAITGP-----AVP-YSRDYKAKYEFERR 309

RESULT 9

T37900

probable ubiquitin-protein ligase - fission yeast (Schizosaccharomyces pombe)

C:Species: Schizosaccharomyces pombe

C>Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 18-Aug-2000

C/Accession: T37900

R:Rieger, M.; McDougall, R.C.; Rajandream, M.A.; Barrell, B.G.

submitted to the EMBL Data Library, September 1999

A:Reference number: Z21752

A/Accession: T37900

A/Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-671 <RIB>

A/Cross-references: EMBL:AL117390; PIDN:CAB55856.1; GSPDB:GN00066; SPDB:SPAC1805.15c

A/Experimental source: strain 972h-; cosmid c1805

C:Genetics:

A:Gene: SPDB:SPAC1805.15c

A:Map position: 1

A:Introns: 60/2; 105/1; 639/2

C:Superfamily: WW repeat homology

F:242-279/Domain: WW repeat homology <WWR>

Query Match 25.9%; Score 1046.5; DB 2; Length 671;

Best Local Similarity 33.9%; Pred. No. 9.8e-66;

Matches 257; Conservative 124; Mismatches 263; Indels 113; Gaps 17;

QY 13 KRLITVLCANLVRKDFRLLPDPFAKVVVDGSGQCHSTDTVNTLDPKWNQHYDLYIGKS 72

Db 7 EVQLTILHVEGLWKNGLRLSLKPYLLISVD-DDQFKITNVASGTLRLSWGFTOKLTVPQ 65

QY 73 DSVTISVWNHKKHKQAGAGLGVRLLSNA---INLKDITGYQRLDCKLGPNDNIVR 129

Db 66 STILLQLQFDEKQ-KNETSDGFVGLGAAVNSFLPFPNPKDDYKTRITL-----RSPSGSYR 120

QY 130 GOIVVSLQSRDRIGTGGQVVD---CSRFPNDLPDGWEERTASGIQVNIHTRITQW 185

Db 121 GS-VVCLFKRSKFLPEELPADKSOICTDIIDASCWETIDFEGHYVL-----170

QY 186 ERPTPEASEYSSPGRPLSCFVDENTPIGNTGATCGGSDPRLAERRVRSORHNYMSRT 245

Db 171 -----KSP-----QLSVISAISHEKLENLTPKQKE--VFSQFLFNNSKS 209

QY 246 HLHTPPD-----LPEGYEQRTTQGGQVYFLHTQTGVSTWHDPRVPRDLNINCEELGPL 299

Db 210 SLKINLEYKVIKHLLEHYPLALSVRQV-----AVEXGFL 244

QY 300 PFGWEIRNTATGRVYFVDHNRRTTQFTDPR-----LSANLHLVLRNQLKDDQQQQQ 351

submitted to the EMBL Data Library, June 1998
A:Description: Arabidopsis thaliana chromosome 1 BAC F1707 sequence.
A:Reference number: Z14334

A:Accession: T01491
A:Status: translated from GB/EMBL/DDBJ
A:Molecule type: DNA

A:Residues: 1-1126 <Y>
A:Cross-references: EMBL:AC003671; NID:g2833627; PID:g3176690; GSPDB:GN00059; ATSP:F1707

A:Experimental source: cultivar Columbia
C:Genetics:
A:Gene: ATSP:F1707.15

A:Map position: 1
A:Introns: 118/3; 224/3; 292/1; 373/3; 457/2; 826/3; 875/3; 902/3; 1057/3; 1093/2

C:Superfamily: ubiquitin-protein ligase homology
F:756-1120/Domain: ubiquitin-protein ligase homology <UBI>

Query Match 19.9%; Score 805.5; DB 2; Length 1126;
Best Local Similarity 41.3%; Pred. No. 6.4e-50;
Matches 177; Conservative 63; Mismatches 142; Indels 47; Gaps 7;

QY 363 LTVPRYKRDVLQKILRQSLQOQPO--ACHCRIEVSREEIPESYRQVQMKRPPKDLWK 420
DB 700 LKAPRL-IDFNKKAYFRSIRHQHDIHSGPLRISVRRAYVLEDSYNQLRMWSPQDLKG 758

QY 421 RLMIKFRGEGLDYGGVAREWLYLLSHEMLNPPYGLFOYSRDDIYTLQINPDSAVNPEHL 480
DB 759 RLNVQFQCEGIDAGGLTREWYQLLSRVIPDKGALLFTTVGND--TFQPNFNSVYQTEHL 817

QY 481 SYHFVGRINGMAVFGHYIDGGFTLPFYKQLLGKSTITLDDMELVDPDLHNSLWILEND 540
DB 818 SYKFGVGRMAKALFDQGLLDVYFTRSFYKHILGVKTYTHDIEAVDPDYKYLKWLLEND 877

QY 541 ITGVLDHTFCVE-----HNAY--GEIQLHKLPGKSIPIVNEENKKEYVRLVYNNWFLRG 593
DB 878 VSDILDITFSMDADEEKHILYKTEVTLVLCMCFPLFFWCFFPKCHCIELIILSLMKKYTDY 937

QY 594 IEAQFLAQGFNEVIPHLLKTFDEKELEILICGLKID----- 633
DB 938 IRPQINAFLEGLNELIPRELVSIFNDKELELLSGLEIDCKLAFYSFLHLFAYSFKV 997

QY 634 -----VNDKVNTRLKHCTPDSNIVKFWKAVFEFDEERRARLLQFVGGSSR 660
DB 998 IITLLSVFFCFLVDLKKANTVTSYTVGSPVIRWVWVKAESKEDMARFLQFVGTGSK 1057

QY 681 VPLOGFKALQAGAPRLFTTHQDACTNNLPKHAHTCFNRIDIPPYSYKLYKLLTAIE 740
DB 1058 VPLEGFKALQISGFORLQHKAYGSPERLPSAHTCFNQLDLPEYOSKEQVQERLLAIH 1117

QY 741 ET---CGFA 746
DB 1118 EANEFGGFA 1126

RESULT 13
H96599
Protein F14716.10 [imported] - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001

R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;

ansen, N.F.; Hughes, B.; Huizart, L.
Nature 408, 816-820, 2000

A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Lueros, J.S.; Maiti, R.; Marziani,

Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,

ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.

A:Reference number: A86141; MUID:21016719; PMID:11130712
A:Status: preliminary
A:Molecule type: DNA

A:Residues: 1-4056 <STO>
A:Cross-references: GB:AE005173; NID:g8778329; PID:AAF79338.1; GSPDB:GN00141

C:Genetics:
A:Gene: F14716.10

A:Map position: 1
Query Match 19.9%; Score 803; DB 2; Length 4056;
Best Local Similarity 40.0%; Pred. No. 6.8e-49;
Matches 177; Conservative 62; Mismatches 143; Indels 60; Gaps 7;

QY 363 LTVPRYKRDVLQKILRQSLQOQPO--ACHCRIEVSREEIPESYRQVQMKRPPKDLWK 420
DB 3617 LKAPRL-IDFNKKAYFRSIRHQHDIHSGPLRISVRRAYVLEDSYNQLRMWSPQDLKG 3675

QY 421 RLMIKFRGEGLDYGGVAREWLYLLSHEMLNPPYGLFOYSRDDIYTLQINPDSAVNPEHL 480
DB 3676 RLNVQFQCEGIDAGGLTREWYQLLSRVIPDKGALLFTTVGND--TFQPNFNSVYQTEHL 3734

QY 481 SYHFVGRINGMAVFGHYIDGGFTLPFYKQLLGKSTITLDDMELVDPDLHNSLWILEND 540
DB 3735 SYKFGVGRMAKALFDQGLLDVYFTRSFYKHILGVKTYTHDIEAVDPDYKYLKWLLEND 3794

QY 541 ITGVLDHTFCVE-----HNAYGEIQLH 562
DB 3795 VSDILDITFSMDADEEKHILYKTEVTLVLCMCFPLFFWCFFPKCHCIELIILSLMKKYTDY 3854

QY 563 ELKPNKSIPIVNEENKKEYVRLVYNNWFLRGIEAQFLAQGFNEVIPHLLKTFDEKEL 622
DB 3855 ELKPGGRNIRVTEETKHEYVDLVAGHILTNAPQINAFLEGFNELIPRELVSIFNDKEL 3914

QY 623 ELICGLKID-----VNDKVNTRLKHCTPDSNIVKFWKAVFEFDEER 667
DB 3915 ELLISGPEIDCKLSDIDQVLCAYPLIDDLKANTVTSYTAGSPVIRWVWVKAESKED 3974

QY 668 RARLLQFVTGSSRVPLOGFKALQAGAPRLFTTHQDACTNNLPKHAHTCFNRIDIPPYS 727
DB 3975 MARFLQFVTGSPVLEGGFKALQISGFORLQHKAYGAPERLPSAHTCFNQLDLPEYQS 4034

QY 728 YEKLYKLLTAIEETC---GFA 746
DB 4035 KEQLQERLLAIHASEGFGFA 4056

RESULT 14
S69625
hypothetical protein YDR457w - yeast (Saccharomyces cerevisiae)

C:Species: Saccharomyces cerevisiae
C:Date: 22-Aug-1996 #sequence_revision 06-Sep-1996 #text_change 23-Mar-2001

C:Accession: S69625
R:Dietrich, F.S.
submitted to the EMBL Data Library, August 1995

A:Description: The sequence of S. cerevisiae cosmid 9410, 8035, 8166, and 9787.
A:Reference number: S69554
A:Accession: S69625

A:Molecule type: DNA
A:Residues: 1-3268 <DIE>
A:Cross-references: EMBL:U33050; NID:g927726; PID:AA864910.1; PID:g927738; MIPS:YDR457w

C:Genetics:
A:Gene: SGD:TOM1
A:Cross-references: SGD:S0002865; MIPS:YDR457w

A:Map position: 4R
Query Match 19.5%; Score 789; DB 2; Length 3268;
Best Local Similarity 45.7%; Pred. No. 5.1e-48;
Matches 164; Conservative 59; Mismatches 132; Indels 4; Gaps 2;

QY 384 SQOQOACHCRIEVSREEIPESYRQVQMKRPPKDLKMLKRIKRGEGLDYGGVAREWL 442
DB 2905 NOERPK---LPITVRRQVFLDSYRALFFKTNDSIKNSKLEITPKGSGVDAGVTREWY 2961

QY 443 YLLSHEMLNPPYGLFOYSRDDIYTLQINPDSAVNPEHLSYFHFVGRINGMAVFGHYIDG 502
DB 2962 QVLSRQMFNPYALFLPVPSDKTTFHNPRTSGINPEHLSFFKFGIMIGKAIHQCCFLDC 3021

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 21, 2004, 07:29:24 ; Search time 17.2889 Seconds
(without alignments)
2252.800 Million cell updates/sec

Title: US-10-009-945-4

Perfect score: 4038

Sequence: 1 MSNFGRRNGPVKRLTLVLC.....EKLVEKLLTAETCTGPAVE 748

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_42:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	4031	99.8	748	1	SUF2 HUMAN
2	3015.5	74.7	757	1	Q9hau4 homo sapien
3	2937.5	72.7	731	1	SUF1_XENLA
4	2491.5	61.7	619	1	SUF1_MOUSE
5	1585.5	35.3	767	1	PBI1_SCHPO
6	1484	36.8	809	1	R3P5_YEAST
7	1358	33.6	887	1	NED4_MOUSE
8	1355.5	33.6	887	1	NED4_RAT
9	1354.5	33.5	1000	1	NED4_HUMAN
10	1336	33.1	864	1	ITCH_MOUSE
11	1313.5	32.5	903	1	ITCH_HUMAN
12	1298.5	32.2	870	1	WPF2_HUMAN
13	1297	32.1	922	1	WPF1_MOUSE
14	1286.5	31.9	870	1	WPF2_MOUSE
15	1281	31.7	918	1	WPF1_MOUSE
16	626	15.7	310	1	URB1_RAT
17	539	13.3	875	1	UR3A_HUMAN
18	484.5	12.0	885	1	UR3A_MOUSE
19	460.5	11.4	892	1	HUL4_YEAST
20	445.5	11.0	1050	1	HUL4_MOUSE
21	404	10.0	1992	1	TRIB_HUMAN
22	342	8.5	910	1	HUL5_YEAST
23	286.5	7.1	1620	1	HED1_HUMAN
24	281	7.0	920	1	EDD_RAT
25	280	6.9	2799	1	EDD_HUMAN
26	278.5	6.9	1483	1	URD4_YEAST
27	262.5	6.5	1647	1	YDE1_SCHPO
28	224.5	5.6	2895	1	HYD_DROME
29	207	5.1	472	1	YAPI_MOUSE
30	161.5	4.0	1277	1	API1_RAT
31	159	3.9	1275	1	API1_MOUSE
32	158.5	3.9	1455	1	API1_HUMAN
33	158	3.9	446	1	YAPI_CHICK

ALIGNMENTS

RESULT 1

ID	SUF2 HUMAN	STANDARD;	PRT;	748 AA.
AC	Q9HAF4; Q9H260;			
DT	28-FEB-2003 (Rel. 41, Created)			
DT	28-FEB-2003 (Rel. 41, Last sequence update)			
DT	10-OCT-2003 (Rel. 42, Last annotation update)			
DE	Smad ubiquitination regulatory factor 2 (EC 6.3.2.-) (Ubiquitin--protein ligase SMURF2) (Smad-specific E3 ubiquitin ligase 2)			
DE	(SMURF2).			
GN	SMURF2.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A., AND MUTAGENESIS OF 251-PRO--VAL-284 AND 297-GLY--LEU-330.			
RX	PubMed=11163210;			
RA	Kavsek P., Rasmussen R.K., Causing C.G., Bonni S., Zhu H.,			
RA	Thomsen G.H., Wrana J.L.;			
RT	"Smad7 binds to Smurf2 to form an E3 ubiquitin ligase that targets the TGF-beta receptor for degradation.";			
RL	Mol. Cell 6:1365-1375(2000).			
RN	[2]			
RP	SEQUENCE FROM N.A., AND MUTAGENESIS OF 251-PRO--VAL-284 AND CYS-716.			
RX	MEDLINE=20538422; PubMed=11016919;			
RA	Lin X., Liang M., Feng X.-H.;			
RT	"Smurf2 is a ubiquitin E3 ligase mediating proteasome-dependent degradation of Smad2 in transforming growth factor-beta signaling.";			
RL	J. Biol. Chem. 275:36818-36822(2000).			
RN	[3]			
RP	SEQUENCE FROM N.A., AND MUTAGENESIS OF CYS-716.			
RX	MEDLINE=21107656; PubMed=1158580;			
RA	Zhang Y., Chang C., Gehling D.J., Hemmati-Brivanlou A., Derynck R.;			
RT	"Regulation of Smad degradation and activity by Smurf2, an E3 ubiquitin ligase.";			
RL	Proc. Natl. Acad. Sci. U.S.A. 98:974-979(2001).			
CC	-I- FUNCTION: Interacts with SMAD1, SMAD2 and SMAD7 in order to trigger their ubiquitination and proteasome-dependent degradation. Enhances the inhibitory activity of SMAD7 and reduces the transcriptional activity of SMAD2. Coexpression of SMURF2 with SMAD1 results in considerable decrease in steady-state level of SMAD1 protein and a smaller decrease of SMAD2 level.			
CC	-I- SUBUNIT: Interacts with SMAD1, SMAD2, SMAD3, SMAD6 and SMAD7 but not SMAD4.			
CC	-I- SUBCELLULAR LOCATION: Nuclear. Cytoplasmic in the presence of SMAD7.			
CC	-I- TISSUE SPECIFICITY: Widely expressed.			
CC	-I- DOMAIN: The second and third WW domains are responsible for interaction with R-SMAD (SMAD1, SMAD2 and SMAD3).			
CC	-I- SIMILARITY: Contains 1 C2 domain.			
CC	-I- SIMILARITY: Contains 3 WW domains.			
CC	-I- SIMILARITY: Contains 1 HECT-type E3 ubiquitin-protein ligase domain.			

34	154.5	3.8	454	1	YAPI_HUMAN	P46937 homo sapien
35	129	3.2	608	1	SAV_DROME	Q9VCR6 drosophila
36	127.5	3.2	1530	1	RIM2_MOUSE	Q9EG77 mus musculus
37	125	3.1	386	1	SAVI_MOUSE	Q8VEB2 mus musculus
38	122.5	3.0	672	1	KPCA_HUMAN	P17252 homo sapien
39	122.5	3.0	672	1	KPCA_RABIT	P10102 oryctolagus
40	122.5	3.0	672	1	KPCA_RAT	P05696 rattus norv
41	122	3.0	637	1	GYPI_YEAST	Q08484 saccharomyc
42	121	3.0	1188	1	RIM2_HUMAN	Q9UQ26 homo sapien
43	120.5	3.0	383	1	SAVI_HUMAN	Q9H4B6 homo sapien
44	120.5	3.0	1813	1	UN13_CAEEL	P27715 caenorhabdi
45	120	3.0	1555	1	RIM2_RAT	Q9J1S1 rattus norv

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; AF310676; AAC45422.1; -;
 CC EMBL; AF301463; AAC25641.1; -;
 CC EMBL; AY014180; AAC50421.1; -;
 CC HSP; Q13526; IPIN.
 CC MM; 605532; -;
 CC
 CC DR GO: 0004842; F:ubiquitin-protein ligase activity; NAS.
 CC DR GO: 0016481; P:negative regulation of transcription; NAS.
 CC DR GO: 0017015; P:regulation of TGFbeta receptor signaling pa. . .; NAS.
 CC DR GO: 0006511; P:ubiquitin-dependent protein catabolism; NAS.
 CC DR InterPro: IPR000008; C2.
 CC DR InterPro: IPR008973; C2_CaLB.
 CC DR InterPro: IPR000569; HECT domain.
 CC DR InterPro: IPR001202; WW_Rep5_WWP.
 CC DR Pfam: PF00168; C2; 1.
 CC DR Pfam: PF00632; HECT; 1.
 CC DR Pfam: PF00337; WW; 3.
 CC DR SMART: SM00239; C2; 1.
 CC DR SMART: SM00119; HECTC; 1.
 CC DR SMART: SM00456; WW; 3.
 CC DR PROSITE; PS00499; C2_DOMAIN_1; 1.
 CC DR PROSITE; PS00004; C2_DOMAIN_2; 1.
 CC DR PROSITE; PS0237; HECT; 1.
 CC DR PROSITE; PS01159; WW_DOMAIN_1; 1.
 CC DR PROSITE; PS00020; WW_DOMAIN_2; 3.
 CC DR Ub1 conjugation pathway; Ligase; Repeat; Nuclear protein.
 CC FT DOMAIN 1 98 C2 DOMAIN.
 CC FT DOMAIN 157 190 WW 1.
 CC FT DOMAIN 251 284 WW 2.
 CC FT DOMAIN 297 330 WW 3.
 CC FT DOMAIN 414 748 HECT.
 CC FT MUTAGEN 251 284 MISSING: ABOLISHES INTERACTION WITH
 CC FT MUTAGEN 297 330 MISSING: ABOLISHES INTERACTION WITH
 CC FT MUTAGEN 716 716 C-2A: LOSS OF ABILITY TO UBIQUITINATE
 CC FT MUTAGEN 716 716 C-2G: LOSS OF ACTIVITY (LOSS OF ABILITY
 CC FT MUTAGEN 716 716 TO UBIQUITINATE SMAD1 AND SMAD2 AND NO
 CC FT MUTAGEN 716 716 DOWN-REGULATION OF SMAD1 AND SMAD2
 CC FT MUTAGEN 716 716 PROTEIN LEVELS).
 CC FT CONFLICT 6 6 G -> R (IN REF. 2).
 CC SQ SEQUENCE 748 AA; 86195 MW; 3042B43A3755762 CRC64;
 CC
 CC Query Match 99.8%; Score 4031; DB 1; Length 748;
 CC Best Local Similarity 99.9%; Pred. No. 6.6e-280;
 CC Matches 747; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 CC
 CC QY 1 MSNPGRRNGFVKLRILVLCANLVKDFRLPDPFAKVVVDGSGGCHSDTDVKNLDPK 60
 CC DB 1 MSNPGRRNGFVKLRILVLCANLVKDFRLPDPFAKVVVDGSGGCHSDTDVKNLDPK 60
 CC
 CC QY 61 WQHYDLYIGKSDVTSVNVNHHKHKQAGFLGCVRLLSNAINLKTGYQRDLCKL 120
 CC DB 61 WQHYDLYIGKSDVTSVNVNHHKHKQAGFLGCVRLLSNAINLKTGYQRDLCKL 120
 CC
 CC QY 121 GPNDNDTVRGQIVVSLQSRDRIGTGGQVDCSRLFNDLPDGEERTASGRQYLNHIT 180
 CC DB 121 GPNDNDTVRGQIVVSLQSRDRIGTGGQVDCSRLFNDLPDGEERTASGRQYLNHIT 180
 CC
 CC QY 181 RTTQWERPTPASYSGRPLSCFVDENTPISTNGATCGQSSDPLAERRVRSQRHN 240
 CC DB 181 RTTQWERPTPASYSGRPLSCFVDENTPISTNGATCGQSSDPLAERRVRSQRHN 240
 CC
 CC QY 241 YMSRTHLTPDLPGEYEQRTTQGGQVYFLHTGTGVSTWHDPRVPRDLSNINCEELGPLP 300
 CC DB 241 YMSRTHLTPDLPGEYEQRTTQGGQVYFLHTGTGVSTWHDPRVPRDLSNINCEELGPLP 300

Db 241 YMSRTHLTPDLPGEYEQRTTQGGQVYFLHTGTGVSTWHDPRVPRDLSNINCEELGPLP 300
 QY 301 PGWEIRNTATGRVYFVDHNNRTTQFTDPRLSANLHLVLRNQNLKDDQQQQVYVSLCPDDT 360
 Db 301 PGWEIRNTATGRVYFVDHNNRTTQFTDPRLSANLHLVLRNQNLKDDQQQQVYVSLCPDDT 360
 QY 361 ECLTVPRYKRDVLVOKLILRQELSSQQQPOQAGHCRIEVSREEIFEESYRQVMKRPKDLWK 420
 Db 361 ECLTVPRYKRDVLVOKLILRQELSSQQQPOQAGHCRIEVSREEIFEESYRQVMKRPKDLWK 420
 QY 421 RLMIKFRGEGLDYGVARWLYLLSHEMLNPPYGLFQYSRDDIYTLQINPDSAVNPEHL 480
 Db 421 RLMIKFRGEGLDYGVARWLYLLSHEMLNPPYGLFQYSRDDIYTLQINPDSAVNPEHL 480
 QY 481 SYFHFVGRINGMAVFHGHYIDGGFTLPFYKQLLGKSTLDDMELVDPDLHNSLVWILEND 540
 Db 481 SYFHFVGRINGMAVFHGHYIDGGFTLPFYKQLLGKSTLDDMELVDPDLHNSLVWILEND 540
 QY 541 ITGVLDHTFCVEHNAYGEIIQHLEKPNKPSIPYNEENKKEYVRLYVNRFLRGIEAQFLA 600
 Db 541 ITGVLDHTFCVEHNAYGEIIQHLEKPNKPSIPYNEENKKEYVRLYVNRFLRGIEAQFLA 600
 QY 601 LQKGFNEVIPHLLKTFDEKELELIICGLGKIDVNDWKVNTLKHCTPDSNIVKWFKAV 660
 Db 601 LQKGFNEVIPHLLKTFDEKELELIICGLGKIDVNDWKVNTLKHCTPDSNIVKWFKAV 660
 QY 661 EFFEERRARLLQFVTSRRVPLQGFALOGAAGPRLFTIHOIDACTNNLPKHAHTCFNRI 720
 Db 661 EFFEERRARLLQFVTSRRVPLQGFALOGAAGPRLFTIHOIDACTNNLPKHAHTCFNRI 720
 QY 721 DIPPESEYKLYEKLTAIBETCGFAVE 748
 Db 721 DIPPESEYKLYEKLTAIBETCGFAVE 748
 RESULT 2
 SUF1_HUMAN
 ID SUF1_HUMAN STANDARD; PRT; 757 AA.
 AC Q9HCE7; O75853; Q9JUT8;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Smad ubiquitination regulatory factor 1 (EC 6.3.2.-) (Ubiquitin--
 DE protein ligase SMURF1) (Smad-specific E3 ubiquitin ligase 1)
 DE (hSMURF1).
 DE SMURF1 OR KIAA1625.
 GN Homo sapiens (Human).
 OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE OF 10-731 FROM N.A. (ISOFORM SHORT).
 RX MEDLINE=9385348; PubMed=10458166;
 RA Zhu H., Kavsak P., Abdollah S., Wrana J.L., Thomsen G.H.;
 RT "A Smad ubiquitin ligase targets the BMP pathway and affects embryonic
 RT pattern formation."
 RL Nature 400:687-693(1999).
 RN [2]
 RP SEQUENCE OF 20-731 FROM N.A. (ISOFORM SHORT).
 RA Stoneking T., Bauer C., O'Neal D.;
 RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A. (ISOFORM LONG).
 RC TISSUE=Brain;
 RX MEDLINE=20450683; PubMed=1097877;
 RA Nagase T., Kikuno R., Nakayama M., Hirose M., Ohara O.;
 RT "Prediction of the coding sequences of unidentified human genes.
 RT XVIII. The complete sequences of 100 new cDNA clones from brain which
 RT code for large proteins in vitro."
 RL DNA Res. 7:273-281(2000).
 CC -!- FUNCTION: Interacts with receptor-regulated SMADs specific for the
 CC bmp pathway (SMAD1 AND SMAD5) in order to trigger their
 CC ubiquitination and degradation and hence their inactivation.

```

CC CC -1- ALTERNATIVE PRODUCTS:
CC CC Event=Alternative splicing; Named isoforms=2;
CC CC Name=Long;
CC CC IsoId=Q9HCE7-1; Sequence=Displayed;
CC CC Name=Short;
CC CC IsoId=Q9HCE7-2; Sequence=VSP_006812;
CC CC -1- SIMILARITY: Contains 1 C2 domain.
CC CC -1- SIMILARITY: Contains 2 WW domains.
CC CC -1- SIMILARITY: Contains 1 HECT-type E3 ubiquitin-protein ligase
CC CC domain.
CC CC -----
CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC CC the European Bioinformatics Institute. There are no restrictions on its
CC CC use by non-profit institutions as long as its content is in no way
CC CC modified and this statement is not removed. Usage by and for commercial
CC CC entities requires a license agreement (See http://www.isb.ch/announce/
CC CC or send an email to license@isb-sib.ch)
CC CC -----
CC CC EMBL; AF193364; AAF08298.2; --
CC CC EMBL; AC004893; AAC82434.1; --
CC CC EMBL; AB046845; BAB13451.1; ALT_INIT.
CC CC HSSP; Q13526; 1PIN.
CC CC MIM; 605568; --
CC CC GO; GO:0005622; C:intracellular; TAS.
CC CC GO; GO:000211; F:protein degradation tagging activity; IDA.
CC CC GO; GO:0004842; F:ubiquitin-protein ligase activity; IDA.
CC CC GO; GO:0030134; P:cell differentiation; IDA.
CC CC GO; GO:0007399; P:ectoderm development; TAS.
CC CC GO; GO:0030514; P:negative regulation of BMP signaling pathway; TAS.
CC CC GO; GO:0006512; P:ubiquitin cycle; IDA.
CC CC InterPro; IPR000008; C2.
CC CC InterPro; IPR008973; C2_CalB.
CC CC InterPro; IPR000569; HECT_domain.
CC CC InterPro; IPR001202; WW_Rp5_WWP.
CC CC Pfam; PF00168; C2; 1.
CC CC Pfam; PF00632; HECT; 1.
CC CC Pfam; PF00397; WW; 2.
CC CC SMART; SM00239; C2; 1.
CC CC SMART; SM00119; HECTC; 1.
CC CC SMART; SM00456; WW; 2.
CC CC PROSITE; PS00499; C2_DOMAIN_1; 1.
CC CC PROSITE; PS00004; C2_DOMAIN_2; 1.
CC CC PROSITE; PS0237; HECT; 1.
CC CC PROSITE; PS01159; WW_DOMAIN_1; 1.
CC CC PROSITE; PS00020; WW_DOMAIN_2; 2.
CC CC KW Ub1 conjugation pathway; Ligase; Repeat; Alternative splicing.
CC CC FT DOMAIN 1 99 C2 DOMAIN.
CC CC FT DOMAIN 234 267 WW 1.
CC CC FT DOMAIN 306 339 WW 2.
CC CC FT DOMAIN 420 757 HECT.
CC CC FT BINDING 725 725 UBIQUITIN.
CC CC FT VARSPIC 269 294 Missing (in isoform Short).
CC CC FT MUTAGEN 725 725 /FTId=VSP_006812.
CC CC SQ SEQUENCE 757 AA; 86113 MW; 89A17FC47B40E9 CRC64;
CC CC -----
CC CC Query Match 74.7%; Score 3015.5; DB 1; Length 757;
CC CC Best Local Similarity 72.2%; Pred. No. 2e-207;
CC CC Matches 574; Conservative 63; Mismatches 73; Indels 85; Gaps 10;
CC CC -----
CC CC 1 MSNPGRRNG-PVKRLTVLCAKLVKDPFRLPDPFAKVVVDGSGQCHSTDVKNLDP 59
CC CC 1 MSNPGTRNGSIRKLTVLCAKLVKDPFRLPDPFAKVVVDGSGQCHSTDVKNLDP 60
CC CC 60 KWNQHYDLVIGKSDSVTVVNNHKKHKQAGFLGCVRLLSNAINRLKDTGYORLDLCK 119
CC CC 61 KWNQHYDLVIGKSDSVTVVNNHKKHKQAGFLGCVRLLSNAINRLKDTGYORLDLCK 120
CC CC 120 LQPNNDNTVRQIVVYLSQRDRIGTGQGVDCSLFNDLPDGWEERTASGRQYLNHI 179
CC CC 121 LNPSTDAVRQIVVYLSQTRDRIGTGGSVDCRGLLENE-----GTWY----- 163

```

```

QY 180 TRTOWERTRPASEYSFGRPLSCFVDENTPISGTNGATCG-----OSSDPLAE 230
Db 164 -----EDSGRPLSCFMEBPAPVTDSTGAAGGCGNCFVESPSODQLQA 209
QY 231 RRVRSORHRNYM-----SRTHLTPDPLPEGYEQRITQOGVYFLHTQTGVSTWHDPRV-- 284
Db 210 QRLRNPDVRGSLQTPQNPHGHQSPBELPEGYEQRITQOGVYFLHTQTGVSTWHDPRIFS 269
QY 285 -----PRDLNINCEELGPIPLPGWEIRNTATGRVYFVDHNN 320
Db 270 PSGTIFGGDAALFYELFLLQGHITSEPRDLNINCEELGPIPLPGWEIRNTATGRVYFVDHNN 329
QY 321 RTTQFTDPLSLNHLVLRNQQLKDDQQQVV-----SLCPDDTECLTVPRYKDLVQKL 376
Db 330 RTTQFTDPLSLNHLVLRNQQLKDDQQQVV-----SLCPDDTECLTVPRYKDLVQKL 382
QY 377 KILRQELSQOQQAQHCRIEVSREEIFESYQVMQRPKDLWKLMIKFRGEGLDYG 436
Db 383 KVLRLHLSLQQAQHCRIEVSREEIFESYQVMQRPKDLWKLMIKFRGEGLDYG 442
QY 437 VAREWLYLLSHMLNDPYGLFYQSRDDIYTLQINPDNANVPHSLSYFHFVGRIMNAVH 496
Db 443 VAREWLYLLSHMLNDPYGLFYQSRDDIYTLQINPDNANVPHSLSYFHFVGRIMNAVH 502
QY 497 GHVIGSGFTLPFYKQLLGSITLDDMELVDPDLHNSLVWILENDITGVLDHFCVHNAY 556
Db 503 GHVINGFTVPFYKQLLGSITLDDMELVDPDLHNSLVWILENDITGVLDHFCVHNAY 562
QY 557 GEIIQHELKPNKSIPIVNEENKKEYVRLVYNNRFRGIEBAQFLALQKGFNEVIPHLLKT 616
Db 563 GRILQHELKPNKSIPIVNEENKKEYVRLVYNNRFRGIEBAQFLALQKGFNEVIPHLLKT 622
QY 617 FDEKELELIICGLKIDVNDKVNRLKHKCTPDSNIVKFWKAVEFFDERBARLLQFVT 676
Db 623 FDEKELELIICGLKIDVNDKVNRLKHKCTPDSNIVKFWKAVEFFDERBARLLQFVT 682
QY 677 GSRVPLQGFKALQ-----GAAGPRLFTIHOIDACTNNLPAKHTCFNRIDIPPYESYEKLYE 733
Db 683 GSTRVPLQGFKALQSGSTGAAGPRLFTIHOIDACTNNLPAKHTCFNRIDIPPYESYEKLYE 742
QY 734 KLLTAIEETCGFAVE 748
Db 743 KLLTAIEETCGFAVE 757
CC -----
CC RESULT 3
CC SUFI_XENLA STANDARD; PRT; 731 AA.
CC ID SUFI_XENLA STANDARD; PRT; 731 AA.
CC AC Q9PUN2;
CC DT 28-FEB-2003 (Rel. 41, Created)
CC DT 28-FEB-2003 (Rel. 41, Last sequence update)
CC DT 28-FEB-2003 (Rel. 41, Last annotation update)
CC DE Smad ubiquitination regulatory factor 1 (EC 6.3.2.-) (Ubiquitin--
CC DE protein ligase SMURF1) (Smad-specific E3 ubiquitin ligase) (XSMURF1).
CC GN SMURF1.
CC OS Xenopus laevis (African clawed frog).
CC OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
CC OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidea; Pipidae;
CC OC Xenopodinae; Xenopus.
CC OC NCBI_TaxID=8355;
CC RN [1]
CC RP SEQUENCE FROM N.A.
CC RC TISSUE=Blascula;
CC RA MEDLINE=99385348; PubMed=10458166;
CC RT Zhu H., Kavsak P., Abdollah S., Wrana J.L., Thomsen G.H.;
CC RT "A SMAD ubiquitin ligase targets the BMP pathway and affects embryonic
CC RT pattern formation.";
CC RL Nature 400:687-693 (1999).
CC CC -1- FUNCTION: INTERACTS WITH RECEPTOR-REGULATED SMADS SPECIFIC FOR THE
CC CC BMP PATHWAY (SMAD1 AND SMAD5) IN ORDER TO TRIGGER THEIR
CC CC UBIQUITINATION AND DEGRADATION AND HENCE THEIR INACTIVATION. MAY
CC CC REGULATE ECTODERMAL DIFFERENTIATION AND PATTERN BY MODULATING BMP
CC CC SIGNALING AND MAY ENHANCE CELLULAR RESPONSIVENESS TO THE SMAD2

```

(ACTIVIN/TGF-BETA) PATHWAY.

CC -1- DEVELOPMENTAL STAGE: EXPRESSED FROM THE EGG STAGE TO THE SWIMMING

CC TADPOLE, WITH MAXIMUM LEVELS OBSERVED IN THE STAGES FROM EGG TO

CC GASTRULA. AT GASTRULATION DISTRIBUTED UNIFORMLY IN EMBRYONIC

CC ECTODERM AND INVOLUTING MESODERM, AND EXPRESSION GRADUALLY

CC LOCALIZES TO THE NERVOUS SYSTEM, AT EARLY TADPOLE STAGES EXPRESSED

CC IN THE CNS, EYE, BRANCHIAL ARCHES, KIDNEY AND SOMITES.

CC -1- SIMILARITY: Contains 1 C2 domain.

CC -1- SIMILARITY: Contains 2 WW domains.

CC -1- SIMILARITY: Contains 1 HECT-type E3 ubiquitin-protein ligase

CC domain.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to license@isb-sib.ch).

CC -----

CC EMBL, AF169310; AD52564.1; --

CC GO: 0005622; C:intracellular; TAS.

CC GO: 0000211; F:protein degradation tagging activity; IDA.

CC GO: 0004842; F:ubiquitin-protein ligase activity; IDA.

CC GO: 00030154; P:cell differentiation; IDA.

CC GO: 0007398; P:ectoderm development; TAS.

CC GO: 0003014; P:negative regulation of BMP signaling pathway; TAS.

CC GO: 0006512; P:ubiquitin cycle; IDA.

CC InterPro: IPR000008; C2.

CC InterPro: IPR008973; C2:CaLB.

CC InterPro: IPR000569; HECT domain.

CC InterPro: IPR001202; WW_Rsp_WWP.

CC Pfam: PF00168; C2: 1.

CC Pfam: PF00632; HECT: 1.

CC Pfam: PF00397; WW: 2.

CC SMART: SM00239; C2: 1.

CC SMART: SM00119; HECT: 1.

CC SMART: SM00456; WW: 2.

CC PROSITE: PS00499; C2 DOMAIN_1; 1.

CC PROSITE: PS00004; C2 DOMAIN_2; 1.

CC PROSITE: PS00237; HECT_1; 1.

CC PROSITE: PS01159; WW DOMAIN_1; 1.

CC PROSITE: PS00020; WW DOMAIN_2; 2.

CC US1 conjugation pathway; Ligase; Repeat.

CC FT DOMAIN 1 99 C2 DOMAIN.

CC FT DOMAIN 233 266 WW 1.

CC FT DOMAIN 279 312 WW 2.

CC FT DOMAIN 394 731 HECT.

CC SEQUENCE 731 AA; 83259 MW; 3CE88E512A42CE2C CRC64;

Query Match 72.7%; Score 2937.5; DB 1; Length 731;

Best Local Similarity 72.7%; Pred. No. 7.1e-202;

Matches 558; Conservative 68; Mismatches 85; Indels 57; Gaps 9;

Qy 1 MSNPGRRNG-PYKRLTVLCANLVKQKFFRLPDPFAKVVDSGSOCHSTDVKNLTDP 59

Db 1 MSNVVTRGSSIRVRVTVICANLAKRDFRLPDPFAKVVDSGSOCHSTDVKNLTDP 50

Qy 60 KWNCHDYLYGKSDSVTISVNNHKKHKGAGFLGCVRLLSNAINRLKDTGYQRLDLCK 119

Db 61 KWNCHDYLYGKSDSVTISVNNHKKHKGAGFLGCVRLLSNAINRLKDTGYQRLDLCK 120

Qy 120 LGFNDNDTVRGQIVSLQSDRGTGQGVVDCSRLEFNDLPDQWEERTASGFIQVLNHI 179

Db 121 LNPTDNDVARGQIVSLQSDRGTGQGVVDCSRLEFNDLPDQWEERTASGFIQVLNHI 162

Qy 180 TRTQWERTRPASEYSSGPRSCFVDENTPIS---GTNGATCGSSDPRLAERVRSQ 236

Db 163 -----LEDTGPRPLSCFMDERAPYTDGPAAGGGPGRLVESPGQEQLQAQ 209

Qy 237 RHRNMYERTHLHTP-----PDLPEGVEQRTQGGVFLHTGTGVSTWHDPRVPRD 287

Db 210 RVRGPEVREHVQTPQNRSHGFSQDLPEGVEQRTQGGVFLHTGTGVSTWHDPRVPRD 269

Qy 288 LSNINCEELGPLPGWEIRNTATGRVYFVDHNNRTTQFTDPRLSANLHLVLNRQNLKQD 347

Db 270 LNSVNCDDGLSLPAGMEVRTTVSGRIYFVDHNNRTTQFTDPR-----LHIIHQSLKSP 325

Qy 348 QQ-----QQVYSLCPDDTECLTVPRYKRDLYQKLKILRQELSQQQPQAGHCRIVSREEIF 403

Db 326 NHAIPVQSGSL--EDGDEPPAQRYERDLVQKLKVLRLHLSLQPOAGHCRIVSREEIF 383

Qy 404 EESYRQWKRPRKDLKRLMKIKRGEGIDYGVAREWLYLLSHEMLNPYGLFOYSRDD 463

Db 384 EESYRQIMKRPKDLKRLMKVFRGEGIDYGVAREWLYLLCHEMLNPYGLFOYSTDN 443

Qy 464 IYTLQINPDGAVNPEHLSYFHFVGRIMGMAVFHGHYIDGGFTLPFYKQLLGKSIITDDME 523

Db 444 IYTLQINPDSINSIPDLHLSYFHFVGRIMGMAVFHGHYINGGFTVPFYKQLLGKSIITDDME 503

Qy 524 LVDPDLHNSLVILENDITGVLDHTPCVHNNAIGETIIOHLEKPKNGKSIIPVNEENKKEYVR 583

Db 504 SVDPDLHNSLVILENDITGVLDHTFCVHNNAIGETIIOHLEKPKNGKSIIPVNEENKKEYVR 563

Qy 584 LYVNRFLRGIEAQFLAQKGFNEVIPHLLKTFDEKELELIICGLKIDVNDKVNTRL 643

Db 564 LYVNRFRMGIEAQFLAQKGFNEVIPHLLKTFDEKELELIICGLKIDVNDKVNTRL 623

Qy 644 KHCTPDSNIVKWPWKAYEFFDEERRARLLOFVTGSSRVPLQGGKALQ---GAAGPRLFTI 700

Db 624 KHCLANSNIQWPFQWAVESFDEERRARLLOFVTGSTRVPLQGGKALQGGSTGAAGPRLFTI 683

Qy 701 HQIDACTNLPKAKHTCFNRDIPYSEYKLYEKLTAETCTGFAVE 748

Db 684 HLIDANTDNPKAKHTCFNRDIPYSEYKLYEKLTAETCTGFAVE 731

RESULT 4

ID SUF1 MOUSE STANDARD; PRT; 619 AA.

AC Q9CUN6;

DT 28-FEB-2003 (Rel. 41, Created)

DT 28-FEB-2003 (Rel. 41, Last sequence update)

DT 15-MAR-2004 (Rel. 43, Last annotation update)

DE Smad ubiquitination regulatory factor 1 (EC 6.3.2.-) (Ubiquitin--

DE protein ligase SMURF1) (Smad-specific E3 ubiquitin ligase 1)

DE (Fragment).

GN SMURF1.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI_TaxID=10090;

RN [1]

SEQUENCE FROM N.A.

RC STRAIN=C57BL/6J; TISSUE=Testis;

RX MEDLINE=22354683; PubMed=12468851;

RA Okazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S.,

RA Nikaide I., Osato R., Saito R., Suzuki H., Yamanaka I., Kiyosawa H.,

RA Yagi K., Tomaru Y., Hasegawa Y., Nogami A., Schonbach C., Gojohori T.,

RA Baldarelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J.,

RA Schriml L.M., Kanapin A., Matsuda H., Batalov S., Beisel K.W.,

RA Blake J.A., Bradt D., Brusic V., Chochia C., Corbani L.E., Cousins S.,

RA Dalla E., Dragani T.A., Fletcher C.F., Forrest A., Frazer K.S.,

RA Gaasterland T., Gariboldi M., Gissi C., Godzik A., Gough J.,

RA Grimmond S., Gustincich S., Hirokawa N., Jackson I.J., Jarvis E.D.,

RA Kanai A., Kawaji H., Kawasawa Y., Kedzierski R.M., King B.L.,

RA Konagaya A., Kurochkin I.V., Lee Y., Lenhard R., Lyons P.A.,

RA Maglott D.R., Maltais L., Marchionni L., McKenzie L., Miki H.,

RA Nagashima T., Numata K., Okido T., Pavan W.J., Pertea G., Resole G.,

RA Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramachandran S.,

RA Ravasi T., Reed J.C., Reed D.J., Reid J., Ring B.Z., Ringwald M.,

RA Sandelin A., Schneider C., Semple C.A., Setou M., Shimada K.,

RA Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita M.,

RA Verardo R., Wagner L., Wabstred C., Wang Y., Watanabe Y., Wells C.,

RA Wilming L.G., Wyshaw-Boris A., Yanagisawa M., Yang L., Yang L.,

Yuan Z., Zavolan M., Zhu Y., Zimmer A., Carninci P., Hayatsu N.,

RA Hirozane-Kishikawa T., Konno H., Nakamura M., Sakazume N., Sato K.,
RA Shiraki T., Waki K., Kawai J., Aizawa K., Arakawa T., Fukuda S.,
RA Hara A., Hashizume W., Imotani K., Ishii Y., Itoh M., Kagawa I.,
RA Miyazaki A., Sakai K., Sasaki D., Shibata K., Shinagawa A.,
RA Yasunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J.,
RA Birney E., Hayashizaki Y.,
RT Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573(2002).
CC -!- FUNCTION: Interacts with receptor-regulated SMADs specific for the
CC bmp pathway (SMAD1 AND SMAD5) in order to trigger their
CC ubiquitination and degradation and hence their inactivation (by
CC similarity).
CC -!- SIMILARITY: Contains 1 C2 domain.
CC -!- SIMILARITY: Contains 2 WW domains.
CC -!- SIMILARITY: Contains 1 HECT-type E3 ubiquitin-protein ligase
CC domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AK015264; BAB29770.2; ALT_INIT.
DR HSP; Q13526; IPIN.
DR MGD; MG1:1923038; 4930431E10Rik.
DR GO; GO:0005622; C:intracellular; ISS.
DR GO; GO:0000211; F:protein degradation tagging activity; ISS.
DR GO; GO:0004842; F:ubiquitin-protein ligase activity; ISS.
DR GO; GO:0030154; P:cell differentiation; ISS.
DR GO; GO:0007398; P:ectoderm development; ISS.
DR GO; GO:0030514; P:negative regulation of BMP signaling pathway; ISS.
DR GO; GO:0006512; P:ubiquitin cycle; ISS.
DR InterPro; IPR000008; C2.
DR InterPro; IPR000569; HECT domain.
DR InterPro; IPR001202; WW_RSP5_WWP.
DR Pfam; PF00632; HECT; 1.
DR Pfam; PF00397; WW; 2.
DR SMART; SM00119; HECTc; 1.
DR SMART; SM00456; WW; 2.
DR PROSITE; PS00499; C2 DOMAIN 1; PARTIAL.
DR PROSITE; PS00004; C2 DOMAIN 2; PARTIAL.
DR PROSITE; PS0237; HECT; 1.
DR PROSITE; PS01159; WW DOMAIN 1; 1.
DR PROSITE; PS00020; WW DOMAIN 2; 2.
KW Ub1 conjugation pathway; Ligase; Repeat.
FT NON_TER 1
FT DOMAIN 125 158 WW 1.
FT DOMAIN 171 204 WW 2.
FT DOMAIN 285 619 HECT.
FT DOMAIN 35 38 POLY-GLY.
SQ SEQUENCE 619 AA; 70902 MW; 4CBE2F8624A7B525 CRC64;

Query Match 61.7%; Score 2491.5; DB 1; Length 619;
Best Local Similarity 72.0%; Pred. No. 3.8e-170;
Matches 471; Conservative 59; Mismatches 75; Indels 49; Gaps 6;

QY 109 DTGYORLDLKGNDNTVGGQIVVLSQSDRDTGCGQVVDGSRFPDNDLPDGNERRT 168
DB 1 DTGYORLDLKNPSDPTDAVRGQIVVLSQSDRDTGCGQVVDGSRFPDNDLPDGNERRT 50

QY 169 ASGRIOYLNHTRTQWERPRPASEYSSPGRLPSCFVDENPTSGTNGATCG----- 221
DB 51 --GTVY-----EDSGPRPISCLNEEPATYDGTGAAGGNCRFV 89

QY 222 --QSSDPLAERRVRSQRHNYM-----SRTHLHTPPDLPPEGYEORTTOGQVYFLHTQTG 275
DB 90 ESPSQNQLLVQRLONEVQGLPQPNRPHGQSPPELPEGYEORTTVOGQVYFLHTQTG 149

QY 276 VSTWHDPRVPRDLSNCEELGPLPWPGEIRNTATGRVYFVDHNNRTTQFDPRLSANLH 335

RESULT 5

PUBL_SCHPO
ID PUBL_SCHPO STANDARD; PRT; 767 AA.
AC Q92462; O14454;
DT 01-NOV-1997 (Rel. 35, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Ubiquitin--protein ligase publi (EC 6.3.2.-).
GN PUBL OR SPAC11G7.02.
OS Schizosaccharomyces pombe (Fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
OX NCBI_TaxID=4896;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96205868; PubMed=8635463;
RA Nefsky B., Beach D.;
RT "Publ acts as an E6-AP-like protein ubiquitin ligase in the
RT degradation of cdc25";
RL EMBO J. 15:1301-1312(1996).
RN [2]
RP SEQUENCE FROM N.A.
RX STRAIN=J27;
RX MEDLINE=97340937; PubMed=9197411;
RA Saleki R., Jia Z., Karagiannis J., Young P.G.;
RT "Tolerance of low pH in Schizosaccharomycetes pombe requires a
RT functioning publ ubiquitin ligase";
RL Mol. Gen. Genet. 254:520-528(1997).
RN [3]
RP SEQUENCE FROM N.A.
RX STRAIN=972;
RX MEDLINE=21848401; PubMed=11859360;
RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
RA Brooks J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,
RA Collins M., Connor R., Cronin A., Davis P., Fellwell T., Fraser A.,
RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,
RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,

Db 150 VSTWHDPRVPRDLSNCEELGPLPWPGEIRNTATGRVYFVDHNNRTTQFDPRLSANLH 205
QY 336 LVLNRNQKLDQKQOQ--QVYSLCPDDTTECLTVPRYKDLVQKILRQELSSQOQPOAGHCR 394
Db 206 HINWOCQLKEPSQPLQNEGSVEDELPAPRYERDLVQKILRQELSSQOQPOAGHCR 265
QY 395 IEVSREEIEFESYRQYVMKRPKDLNWKRLMKFGEGBGLDYGGVAREWLXLLSHEMLNPPY 454
Db 266 IEVSREEIEFESYRQYVMKRPKDLNWKRLMKFGEGBGLDYGGVAREWLXLLSHEMLNPPY 395
QY 455 GLFOYRSDDIYTLQINPDNAVNPSHLSYHFVGRINGMAVFGHYIDGGFTLPFYKQLLG 514
Db 326 GLFOYSTDNITLQINPDNSINPDHLSYHFVGRINGMAVFGHYIDGGFTLPFYKQLLG 385
QY 515 KSITLDDMELVDPDLHNSLVILENDITGVLDHTFCVHNAYGEIIQHELKPKNGKIPVN 574
Db 386 KPIQLSDLESVDPELKHSLVILENDITPVLDHTFCVHNAYGEIIQHELKPKNGRNPVT 445
QY 575 EENKKEYVRLYVNWRFRLGIEAQLQKGFNEVIFQHLKTPDEKELELIICGLSKIDV 634
Db 446 EENKKEYVRLYVNWRFRLGIEAQLQKGFNEVIFQHLKTPDEKELELIICGLSKIDV 505
QY 635 NDWKVNTRLKHCTPDNSNIVKWFKAVEFDEERRARLLQFVTGSSRVPLQGGFKALQGAAG 694
Db 506 NDWKSNTRLKHCVADSNIVKWFQAVETDEERRARLLQFVTGSTRVPLQGGFKALQGAAG 565
QY 695 PRFLTTHQIDACTNNLPKHAHTCFNRIDIPPEYSEYKLEKLTATETCGFAVE 748
Db 566 PRFLTTHLIDANTDNLPAHTCFNRIDIPPEYSEYKLEKLTATETCGFAVE 619

QY	322	TTQFTDPRLSANLHLVLRNQRLKQDQOQQVVSLCPDDTECLTVPR--YKRDVLQVKIL	379
DB	598	RTQWEDPRL-----ENVAITGPAVYSYDKRKRYEFF	629
QY	380	ROELSOQQPQAGHCRIEYSREBIFESYQVQMKRPKDLWK-RLMIKFRGEGLDYGGA	438
DB	630	RRKLKQNDIPNKFEMKLRATVLEDSYRRINGVKRADFLKARLWIEPDGKGLDYGGA	689
QY	439	REWLKLSHEMLNPPYGLFQYSRDDIYTLQINPDSAV-NPEHLISYFHFVGRIMGMAVPHG	497
DB	690	REWFLLISKEMFNPPYGLFYSATDNYTLQINPNSGLCNEHLISYFHFVGRIMGMAVPHG	749
QY	498	HYIDGGFTLPPFYKOLLGKSIITLDDMELVDPDLHSLVWLINENDITGLVDHTFCVHEHAYG	557
DB	750	KULDGFPIRPFYKMLKRPITLHDMESVDSEYNSLRWLINENDT-ELDLRFIDDEBLFG	808
QY	558	EIQHELKPNKXSIPVNEENKKEYVRLVYNNRFRUGIEAQFLAQKQGNVPIPHLLKTF	617
DB	809	QTHQHELKNGSSEIVVTNKNKKEYIYLVQWRFVNRIRIQKQMAAFKGGFFELIPQDLIKIF	868
QY	618	DEKELELIICGLGKIDVNDKVNTELKH-CTPDSNIVKWKVAFVFEDEERRARLLQFVT	676
DB	869	DENELLMCGLGVDVNDVNDREHTYKNGYSANHQVQWFKVAMVMDSEKRIQLQFVT	928
QY	677	GSRVPLQGFALQGAAGPRLFTIHQIDACTNNLPKATCFNCRNIDIPPEYSEKLYEKIL	736
DB	929	GTSRVPMNGFAELGSGNGPQSFTVEQW-GTPEKLPRAHTCFNRLDLPPESFEELWDLKQ	987
QY	737	TALETCTGF 745	
DB	988	MAIENTQGF 996	
RESULT 10			
ITCH_MOUSE			
ID	ITCH_MOUSE	STANDARD;	PRT; 864 AA.
AC	Q8C863; O54971;		
DT	10-OCT-2003 (Rel. 42, Created)		
DT	10-OCT-2003 (Rel. 42, Last sequence update)		
DT	15-MAR-2004 (Rel. 43, Last annotation update)		
DE	Itchy E3 ubiquitin protein ligase (EC 6.3.2.-)		
GN	ITCH.		
OS	Mus musculus (Mouse).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
OX	NCBI_TaxID=10090;		
RN	[1]		
RP	SEQUENCE FROM N.A. (ISOFORM 1), AND DISEASE.		
RC	STRAIN=C3H/HeJ; TISSUE=Kidney;		
RX	MEDLINE=9812574; PubMed=9462742;		
RA	Perry W.L., Huetad C.M., Swing D.A., O'Sullivan T.N., Jenkins N.A.,		
RA	Copeland N.G.		
RT	"The itchy locus encodes a novel ubiquitin protein ligase that is		
RL	disrupted in a18H mice."		
RL	Nat. Genet. 18:143-146(1998).		
RN	[2]		
RP	SEQUENCE FROM N.A. (ISOFORM 2).		
RC	STRAIN=C57BL/6J; TISSUE=Head;		
RX	MEDLINE=22354683; PubMed=12466851;		
RA	Okazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S.,		
RA	Nikaido I., Oesato N., Saito R., Suzuki H., Yamanaka I., Kiyosawa H.,		
RA	Yagi K., Tomaru Y., Hasegawa Y., Nogami A., Schonbach C., Gojobori T.,		
RA	Baldarelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J.,		
RA	Shiremali L.M., Kanapin A., Matsuda H., Batalov S., Beisel K.W.,		
RA	Blake J.A., Braggi T.A., Fletcher C.F., Forrest A., Frazer K.S.,		
RA	Dalla E., Dragani T.A., Fletcher C.F., Godzik A., Gough J.,		
RA	Gaasterland T., Gariboldi M., Gissi C., Godzik A., Gough J.,		
RA	Grimmond S., Gustincich S., Hirokawa N., Jackson I.J., Jarvis E.D.,		
RA	Kanai A., Kawaji H., Kawasawa Y., Kedzierski R.M., King B.L.,		
RA	Kanagaya A., Kurochkin I.V., Lee Y., Lennard B., Lyons P.A.,		
RA	Maglott D.R., Maltais L., Marchionni L., McKenzie L., Miki H.,		
RA	Nagashima T., Numata K., Okido T., Pavan W.J., Pertea G., Pesole G.,		
RA	Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramachandran S.,		

RA Ravasi T., Reed J.C., Reed D.J., Reid J., Ring B.Z., Ringwald M.,
RA Sandelin A., Schneider C., Sempke C.A., Setou M., Shimada K.,
RA Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita M.,
RA Verardo R., Wagner L., Wahlestedt C., Wang Y., Watanabe Y., Wells C.,
RA Wilming L.G., Wynshaw-Boris A., Yanagisawa M., Yang I., Yang L.,
RA Yuan Z., Zavolan M., Zhu Y., Zimmer A., Carninci P., Hayatsu N.,
RA Hirozane-Kishikawa T., Konno H., Nakamura M., Sakazume N., Sato K.,
RA Shiraki T., Waki K., Kawai J., Aizawa K., Arakawa T., Fukuda S.,
RA Hara A., Hashizume W., Imotani K., Ishii Y., Itoh M., Kagawa I.,
RA Miyazaki A., Sakai K., Sasaki D., Shibata K., Shinagawa A.,
RA Yasunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J.,
RA Birney E., Hayashizaki Y.,
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs";
RL Nature 420:563-573 (2002).
RN [3]
RN FUNCTION, AND INTERACTION WITH JUN AND JUNB.
RP MEDLINE=21864584; PubMed=11828324;
RX Fang D., Elly C., Gao B., Fang N., Altman Y., Joazeiro C., Hunter T.,
RA Copeland N., Jenkins N., Liu Y.C.;
RA "Dysregulation of T lymphocyte function in itchy mice: a role for
RT Itch in TH2 differentiation";
RT Nat. Immunol. 3:281-287 (2002).
RN [4]
RN INTERACTION WITH NOTCH1, AND MUTAGENESIS OF CVS-832.
RP MEDLINE=20549573; PubMed=10940313;
RX Qiu L., Joazeiro C., Fang N., Wang H.-Y., Elly C., Altman Y., Fang D.,
RA Hunter T., Liu Y.-C.;
RA "Recognition and ubiquitination of Notch by Itch, a hec-type E3
RT ubiquitin ligase";
RL J. Biol. Chem. 275:35734-35737 (2000).
RN [5]
RN INTERACTION WITH OGNL.
RX MEDLINE=21895815; PubMed=11782481;
RA Traweger A., Fang D., Liu Y.-C., Stelzhammer W., Krizbai I.A.,
RA Fresser F., Bauer H.-C., Bauer H.;
RA "The tight junction-specific protein occludin is a functional target
RT of the E3 ubiquitin-protein ligase itch";
RL J. Biol. Chem. 277:10201-10208 (2002).
CC -!- FUNCTION: E3 ubiquitin-protein ligase which accepts ubiquitin from
CC an E2 ubiquitin-conjugating enzyme in the form of a thioester and
CC then directly transfers the ubiquitin to targeted substrates.
CC Regulates the transcriptional activity of several transcription
CC factors, and probably plays an important role in the regulation of
CC immune response.
CC -!- PATHWAY: Ubiquitin conjugation; third step.
CC -!- SUBUNIT: Interacts via its WW domains with OCLN, NOTCH1, JUN and
CC JUNB. Interacts with DRP1A and NFE2 (By similarity).
CC -!- SUBCELLULAR LOCATION: Nuclear and cytoplasmic.
CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=2;
CC Name=1;
CC IsoId=Q8C863-1; Sequence=Displayed;
CC Note=Major form;
CC Name=2;
CC IsoId=Q8C863-2; Sequence=VSP_008452, VSP_008453;
CC -!- TISSUE SPECIFICITY: Widely expressed.
CC -!- PTM: Phosphorylated on tyrosine residues (By similarity).
CC -!- DISEASE: Defects in Itch are the cause of the itchy phenotype
CC which is an inflammatory and immunological condition characterized
CC by inflammation in the lung and stomach, hyperplasia in lymphoid
CC and hematopoietic cells and constant itching in the skin.
CC -!- SIMILARITY: Contains 1 C2 domain.
CC -!- SIMILARITY: Contains 1 HECT-type E3 ubiquitin-protein ligase
CC domain.
CC -!- SIMILARITY: Contains 4 WW domains.
CC -!- CAUTION: It is uncertain whether Met-1 or Met-11 is the initiator.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AF037454; AAB99764.1; ALT_INIT.
CC EMBL; AK048303; BAC33298.1; -.
DR HSSP; Q13526; 1PIN.
DR MGD; MGI:1202301; Itch.
DR InterPro; IPR000008; C2.
DR InterPro; IPR008973; C2_CaLB.
DR InterPro; IPR000569; HECT_domain.
DR InterPro; IPR002349; WW.
DR InterPro; IPR001202; WW_Rsp5_WWP.
DR Pfam; PF00168; C2; 1.
DR Pfam; PF00632; HECT; 1.
DR Pfam; PF00397; WW; 4.
DR PRINTS; PRO0403; WWDOMAIN.
DR SMART; SM00239; C2; 1.
DR SMART; SM00119; HECTC; 1.
DR SMART; SM00456; WW; 4.
DR PROSITE; PS00499; C2_DOMAIN_1; FALSE_NEG.
DR PROSITE; PS00004; C2_DOMAIN_2; 1.
DR PROSITE; PS50237; HECT; 1.
DR PROSITE; PS01159; WW_DOMAIN_1; 4.
DR PROSITE; PS50020; WW_DOMAIN_2; 4.
DR Ubl conjugation pathway; Ligase; Nuclear protein; Repeat;
KW Phosphorylation; Alternative C2 DOMAIN.
FT DOMAIN 5 99
FT DOVAIN 287 320 WW 1.
FT DOVAIN 319 352 WW 2.
FT DOVAIN 399 432 WW 3.
FT DOVAIN 439 472 WW 4.
FT DOVAIN 530 864 HECT.
FT DOVAIN 218 223 POLY-PRO.
FT VARSPLIC 742 759 LLCQMQLDNDWQWRAI -> MNFYLLKHTSKYSRYLF
FT (in isoform 2).
FT VARSPLIC 760 864 Missing (in isoform 2).
FT MUTAGEN 832 832 /FTID=VSP_008453.
FT C->A: LOSS OF UBIQUITIN PROTEIN LIGASE
FT ACTIVITY.
SQ SEQUENCE 864 AA; 98993 MW; 905FDBE0A1EA7EA CRC64;
Query Match 33.1%; Score 1336; DB 1; Length 864;
Best Local Similarity 36.6%; Pred.No. 1.8e-87;
Matches 316; Conservative 130; Mismatches 272; Indels 146; Gaps 24;
QY 13 KRLTTLVCAK-NLVKQDFFRLDPFAKVVWDGSGQCHSTDTKNTLDPKWNQHYDLYIGK 71
DB 19 QLQITVISAKLKENKQWFG-PSYVEVTVD--GQSKTEKCNNTNSPKWKQPLTVIVTP 75
QY 72 SDSVTISVNNHKKHK--QAGFLGCVRLLSNAINRLKDTGYQRLDLCKLGPNDNVTVR 129
DB 76 TSKLCFRVWSHQTLKSDVLLTAGLDIYETLKSNMKLEEV---VMTLQLVGDKPEPTETM 132
QY 130 GOIVV---SLOSRRDRTGTGGGVV-----DCSRLFDNDLPD-----GWEE 165
DB 133 GDLSCVLDGLQVEAEVVTNGTSCSESTQDDQCRTRDTRVSTNGSEPEVAASENK 192
QY 166 RRTASGRITQYLNHTRITQWERTPRPASEVSSGPRLSCFVDENT-----PIS 213
DB 193 RANGNNSPLSNGGFKFSRPPRPPPP--PTPRPASVNGSPSTNSDGSSTGSLPPT 250
QY 214 GTN-----CATCG-----QSSDPLAERRVRSRHRNYM--- 242
DB 251 NTNVTSTEGATSGLIPLITISGSGRPLNTVSQLPPGMEQV--DQGRVYVDHV 309
QY 243 -SRTHLTPPLPSPGYEQRTTQQGVYFLHTQTGVSTWHDPRVP----- 285
DB 310 EKRTTWDPRPELPGWERRVDNMGRIYVDHFTTTTQWRPTLESVRNVEQWQQLRSQ 369
QY 286 -----RDL-----SNINCEELGPLPGWEIRNTATGEVYVDHNNETTFD 327
DB 370 GMAQOQNFRIYGNQDLFATSQNKFEPLGLPGLPGWEKRTDSNGRVYFVNHTITQWED 429

QY 328 PRLSANLHLVLRNQLKQDQOQVVSCLP---DTECLTV---PR-----Y 368
 Db 430 PRSQOQ---LNEKLPFGWEHRTVDGIPYFVDHNRATTVIDPTGKSLDNGPQIAY 485
 QY 369 KRDVLQKILR---QELSQOQFQAGHCRIEVSREEIFEESYRQVMKVRPKDLKRLMIK 425
 Db 486 VRDFRAKVQYFRFCQQLAMPO---HIKITVTRTLFEDSFQIQIMSFQDRLRLWVI 541
 QY 426 FGEGLDGGVAREWYLHSHMLNPPYGLFQYSRDDIYTLQINPDVAVNPEHLSYHPF 485
 Db 542 FGEGLDGGVAREWYFLSHSEVLNPMYCLFEYAGKNYCLQINPASYINPDHLKYRPF 601
 QY 486 VGRIMGVAFVGHYIDGGFTLFFYKQLLCKSTLDDMELVDPDLHNSLVWILENDITGV- 544
 Db 602 IGRFIAMALFHGKFDITGSLFFYKILNPKVGLKDLSDIDPEFNSLWYKNNISECG 661
 QY 545 LDHTCQVEINAYGEIICHHLKPKNGKSIPIVNBENKKEYVRLYNWFLRGIEAQFIALQKG 604
 Db 662 LEMYSVDKEILGEIKHDLKPKNGNGLVTEENKEEYIRMAVMAELSRGVEEQTAFFEG 721
 QY 605 FNEVIPQHLKTFDEKELELIICGLGKIDVNDKYNTRLKHCTPDSTNVKFWKAVEFFD 664
 Db 722 FNEIIPQVLYQVDAKELEVLGCGHQLDNDQWQHAIRYHRTSRKQIMFWQVKEID 781
 QY 665 EERRARLQFVTGSSRVPVLPQGFALQGAAGPLFTIHQIDACTNNLPKAHTCFNRDIPP 724
 Db 782 NEKMRLLQFVTGTCRLPVGGFADLMGNGPQKFCIEKVGK-ENWLPRSHTCFNRDLDP 840
 QY 725 YESYEKLYEKLATATEETCGRAVE 748
 Db 841 YKSYBQLEKLLFAIETEETGFGQE 864

RESULT 11
 ITCH HUMAN
 ID ITCH HUMAN STANDARD; PRT: 903 AA.
 AC Q96J02; O43584; Q96F66; Q9BV75; Q9H451; Q9H4U5;
 DT 10-OCT-2003 (Rel. 42, Created)
 DT 10-OCT-2003 (Rel. 42, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Itchy homolog E3 ubiquitin protein ligase (EC 6.3.2.-) (ITCH)
 DE (Atrophin-1-interacting protein 4) (AIP4) (NFE2-associated polypeptide
 DE 1) (NAPP1).
 GN ITCH.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A. (ISOFORM 2), AND INTERACTION WITH NFE2.
 RC TISSUE=Leukemia;
 RX MEDLINE=21218930; PubMed=11318614;
 RA Chen X., Wen S.-C., Fukuda M.N., Gavva N.R., Hsu D.-W., Akama T.O.,
 RA Yang-Peng T.L., Shen C.K.J.;
 RT "Human ITCH is a co-regulator of the hematopoietic transcription
 RT factor NF-E2".
 RL Genomics 73:238-241(2001).
 RN [2]
 RP SEQUENCE FROM N.A. (ISOFORM 2).
 RA Miyazaki K., Okamoto Y., Sakamoto M., Nakagawara A.;
 RT "Homo sapiens mRNA for ubiquitin protein ligase Itch, complete cds.";
 RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21638749; PubMed=11780052;
 RA Deloukas P., Matthews L.H., Ashurst J., Burton J., Gilbert J.G.R.,
 RA Jones M., Stavrides G., Almeida J.P., Babbage A.K., Bagguley C.L.,
 RA Bailey J., Barlow K.F., Bates K.N., Beard L.M., Beare D.M.,
 RA Beasley O.P., Bird C.P., Blakey S.E., Bridgeman A.M., Brown A.J.,
 RA Buck D., Burrill W., Butler A.P., Carder C., Carter N.P., Clee C.M.,
 RA Chapman J.C., Clamp M., Clark L.N., Clark S.Y., Clegg S., Clegg S.,
 Clegg S., Cogley V.E., Collier R.E., Connor R., Corby N.R.,

RA Coulson A., Coville G.J., Deadman R., Dhani P., Dunn M.,
 RA Ellington A.G., Frankland J.A., Fraser A., French L., Garner P.,
 RA Graffham D.V., Griffiths C., Griffiths M.N.D., Gwilliam R., Hall R.E.,
 RA Hammond S., Harley J.L., Heath P.D., Ho S., Holden J.L., Howden P.J.,
 RA Huckle E., Hunt A.R., Hunt S.E., Jekosch K., Johnson C.M., Johnson D.,
 RA Kay M.P., Kimberley A.M., King A., Knights A., Laird G.K., Lawlor S.,
 RA Levasialho M.H., Leversha M., Lloyd C., Lloyd D.M., Lovell J.D.,
 RA Marsh V.L., Martin S.L., McConnachie L.J., McElay K., McMurray A.A.,
 RA Milne S., Mistry D., Moore M.J.F., Mullikin J.C., Nickerson T.,
 RA Oliver K., Parker A., Patel R., Pearce T.A.V., Peck A.I.,
 RA Phillimore B.J.C.T., Prathalingam S.R., Plumb R.W., Ramsey H.,
 RA Rice C.D., Ross M.T., Scott C.E., Sehra H.K., Showkhen R., Sims S.,
 RA Skuce C.D., Smith M.L., Soderlund C., Steward C.A., Sulston J.E.,
 RA Swann A., Sycamore N., Taylor R., Tee L., Thomas D.W., Thorpe A.,
 RA Tracey A., Tromans A.C., Vaudin M., Wall M., Wallis J.M.,
 RA Whitehead S.L., Whittaker P., Willey D.L., Williams L., Williams S.A.,
 RA Wilming L., Wray P.W., Hubbard T., Durbin R.M., Bentley D.R., Beck S.,
 RA Rogers J.;
 RL "The DNA sequence and comparative analysis of human chromosome 20.";
 RL Nature 414:865-871(2001).
 RN [4]
 RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).
 RC TISSUE=Kidney, and Placenta;
 RX MEDLINE=22388257; PubMed=12477932;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaudo M.F., Casavant T.B., Scheetz T.E.,
 RA Brownstein M.J., Ustin T.B., Toshikiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Kettelman M., Maman A., Rodriguez S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalish D.E.,
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 RT and mouse cDNA sequences".
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16999-16903(2002).
 RN [5]
 RP SEQUENCE OF 83-903 FROM N.A. (ISOFORM 2), AND INTERACTION WITH DRPLA.
 RC TISSUE=Fetal brain;
 RX MEDLINE=98313405; PubMed=9647693;
 RA Wood J.D., Yuan J., Margolis R.L., Colomer V., Duan K., Kushi J.,
 RA Kaminsky Z., Kleiderlein J.J., Sharp A.H., Ross C.A.;
 RT "Atrophin-1, the DRPLA gene product, interacts with two families of WW
 RT domain-containing proteins".
 RL Mol. Cell. Neurosci. 11:149-160(1998).
 RN [6]
 RP SEQUENCE OF 463-470; 503-510; 514-526; 644-665 AND 875-881,
 RP INTERACTION WITH LMP2A, AND MUTAGENESIS OF CYS-871.
 RC TISSUE=B-cell;
 RX MEDLINE=20501262; PubMed=11046148;
 RA Winberg G., Matekova L., Chen F., Plant P., Rotin D., Gish G.,
 RA Ingham R., Ernberg I., Pawson T.;
 RT "Latent membrane protein 2A of Epstein-Barr virus binds WW domain E3
 RT protein-ubiquitin ligases that ubiquitinate B-cell tyrosine kinases";
 RL Mol. Cell. Biol. 20:8526-8535(2000).
 RN [7]
 RP INTERACTION WITH CBLC, AND PHOSPHORYLATION.
 RX MEDLINE=22323223; PubMed=1226085;
 RA Courbard J.-R., Fiore F., Adelaide J., Borg J.P., Birnbaum D.,
 RA Ollendorff V.;
 RT "Interaction between two ubiquitin-protein isopeptide ligases of
 RT different classes, CBLC and AIP4/ITCH".
 RL J. Biol. Chem. 277:45267-45275(2002).
 CC -!- FUNCTION: E3 ubiquitin-protein ligase which accepts ubiquitin from
 CC an E2 ubiquitin-conjugating enzyme in the form of a thioester and

[1]
 RP SEQUENCE FROM N.A., AND INTERACTION WITH WBP1; WBP2; SCNN1A; SCNN1B
 RP AND SCNN1G.
 RC TISSUE=Bone marrow, and Brain;
 RX MEDLINE=22388257; PubMed=12477932;
 RA Klausner R.L., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.B.,
 RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Kettner M., Madan A., Rodrigues S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickens M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butterfield V.S.N., Krzywinski M.I., Skalska U., Smalls D.E.,
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length
 human and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
 RN [3]
 RP INTERACTION WITH DRPLA, AND TISSUE SPECIFICITY.
 RX MEDLINE=98313405; PubMed=9647693;
 RA Wood J.D., Yuan J., Margolis R.L., Colomer V., Duan K., Kushi J.,
 RA Kaminsky Z., Kleiderlein J.J. Jr., Sharp A.H., Ross C.A.;
 RT "Atrophin-1, the DRPLA gene product, interacts with two families of
 RT WW domain-containing proteins.";
 RL Mol. Cell. Neurosci. 11:149-160 (1998).
 RN [4]
 RP INTERACTION WITH SCNN1A; SCNN1B AND SCNN1G.
 RX MEDLINE=22157111; PubMed=12167593;
 RA McDonald F.J., Western A.H., McNeil J.D., Thomas B.C., Olson D.R.,
 RA Snyder P.M.;
 RT "Ubiquitin-protein ligase WBP2 binds to and downregulates the
 RT epithelial Na(+) channel.";
 RL Am. J. Physiol. 283:F431-F436 (2002).
 RN [5]
 RP INTERACTION WITH ADENOVIRUS TYPE 2 PIII.
 RX MEDLINE=22338357; PubMed=12450395;
 RA Gallinier R., Gout E., Lortat-Jacob H., Wood J., Chroboczek J.;
 RT "Adenovirus protein involved in virus internalization recruits
 RT ubiquitin-protein ligases.";
 RL Biochemistry 41:14298-14305 (2002).
 CC -!- FUNCTION: E3 ubiquitin-protein ligase which accepts
 CC ubiquitin from an E2 ubiquitin-conjugating enzyme in the form of a
 CC thioester and then directly transfers the ubiquitin to targeted
 CC substrates (By similarity).
 CC -!- SUBUNIT: Ubiquitin conjugation; third step.
 CC -!- SUBUNIT: Binds SCNN1A, SCNN1B, SCNN1G, WBP1, WBP2, DRPLA and
 CC adenovirus type 2 PIII.
 CC -!- TISSUE SPECIFICITY: Detected in heart, throughout the brain,
 CC placenta, lung, liver, muscle, kidney and pancreas.
 CC -!- MISCELLANEOUS: A cysteine residue is required for ubiquitin-
 CC thiolester formation.
 CC -!- SIMILARITY: Contains 1 C2 domain.
 CC -!- SIMILARITY: Contains 4 WW domains.
 CC -!- SIMILARITY: Contains 1 HECT-type E3 ubiquitin-protein ligase
 CC domain.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; U96114; AAC51325.1; -;
 CC DR EMBL; BC000108; AAH00108.1; -;
 CC DR EMBL; BC013645; AAH13645.1; -;
 CC DR HSSP; Q13526; IPIN.
 CC DR MIM; 602308; -;
 CC DR GO; GO:0000151; C:ubiquitin ligase complex; TAS.
 CC DR GO; GO:0004942; F:ubiquitin-protein ligase activity; TAS.
 CC DR GO; GO:0006464; P:protein modification; TAS.
 CC DR InterPro; IPR000008; C2_CaLB.
 CC DR InterPro; IPR008973; C2_CaLB.
 CC DR InterPro; IPR000569; HECT_domain.
 CC DR InterPro; IPR02349; WW.
 CC DR Pfam; PF00168; C2; 1.
 CC DR Pfam; PF00632; HECT; 1.
 CC DR Pfam; PF00397; WW; 4.
 CC DR PRINTS; PR00403; WWDOMAIN.
 CC DR SMART; SM00239; C2; 1.
 CC DR SMART; SM00119; HECTC; 1.
 CC DR SMART; SM00456; WW; 4.
 CC DR PROSITE; PS00499; C2_DOMAIN_1; FALSE_NEG.
 CC DR PROSITE; PS00004; C2_DOMAIN_2; FALSE_NEG.
 CC DR PROSITE; PS0237; HECT; 1.
 CC DR PROSITE; PS01159; WW_DOMAIN_1; 4.
 CC DR PROSITE; PS00020; WW_DOMAIN_2; 4.
 CC DR PROSITE; PS00020; WW_DOMAIN_2; 4.
 CC KW Ub1 conjugation pathway; Ligase; Repeat.
 CC FT DOMAIN 20 100 C2 DOMAIN.
 CC FT DOMAIN 300 333 WW 1.
 CC FT DOMAIN 330 363 WW 2.
 CC FT DOMAIN 405 437 WW 3.
 CC FT DOMAIN 444 477 WW 4.
 CC FT DOMAIN 536 870 HECT.
 CC FT BINDING 838 838 UBIQUITIN (BY SIMILARITY).
 CC FT BINDING 136 136 E -> K (IN REF. 1).
 CC FT CONFLICT 394 395 SS -> FW (IN REF. 1).
 CC SQ SEQUENCE 870 AA; 98911 MW; FCCD75CBA61F2204 CRC64;
 CC
 CC Query Match 32.2%; Score 1298.5; DB 1; Length 870;
 CC Best Local Similarity 42.0%; Pred. No. 8.7e-85;
 CC Matches 27; Conservative 100; Mismatches 217; Indels 57; Gaps 9;
 CC
 CC QY 122 PNNDTVRGQIVVLSQSRDRIGTGGOVDCSRFLDNDLPDGWEERTASGRIOYLNHITR 191
 CC DB PNTTSLPAPATPAEGEPSTSGTQQLPAAQAPDAPAGWEORELPNGRVVYVDHTK 324
 CC QY 182 TTOWERPTRPASYSFGRPLSCFVDENTPI SGTNATCGQSDPPLAERRVRSQRHNY 241
 CC DB TTTWERPLPGWEKTRDPRGRFYVDHNT-----RTTTWQRTAEYVRYNEQWQSRNQ 378
 CC QY 242 MSKTHLH-----TPED-----LPBGVEQRTTQGOVYFLHTGTGSTDHPRV 284
 CC DB LQAGAMHFSORFLYQSSASTDHPGLPPGWEKR-QDNGRVVYVNNHNTTRTTQWEDPT 437
 CC QY 285 PRDLSNINCEELGPPPGWEIRNTATGRVYFVDHNNRTQFTDPRLSANLHLVLRNQNL 344
 CC DB QGMIQE-----PALPGWEMKYTSGVYFVDHNTTRTTTFKDPFGFE----- 480
 CC QY 345 KDQQQQVSLCPDDTECLTPYKEDLVQKLILRQELSQQQPQAGHCIEVSRREIFE 404
 CC DB -----SGTKQSGPGAYDRSFRWKYQFPR-FLCHSNALPSHKVKSIVSRQTLFE 526
 CC QY 405 ESYRQVMKMRPKDLWKRLMKFRGEGLDYGGVAREWLYLLSHEMLNPPYGLFOYGRDDI 464
 CC DB DSFQQINWKPDLRLRLYIIMRGEGLDYGGIAREWFFLLSHEVLNPNVCLFEYAGKNN 586

QY 465 YTLQINPDSAVNPEHLSYFHFVGRINGMAVFHGHYIDGGFTLFPYKQLLGKSTLDDMEL 524
DB 587 YCQINPASSINDHLTYFRIGRFAMADYHGKFDITGTLFPYKMLNKPTLKDLES 646
QY 525 VDPDLHNSLVILENDITGV-LDHTFCVEHAYGEIIOHELKPNKSGSIPVNEENKEYVR 583
DB 647 IDPEFYNSIWMKNNLECEGLYFIQDMELIGKVTHELKEGGSIRVTENKEEYIM 706
QY 584 LYVNRRLRIGIEAFLAQGRNEV:POHLLKTDFEKELELLICGLKIDVNDKNTL 643
DB 707 LLDWRETRGVBEQTKAFLDGFNEVAPLEWLRVDFEKELEMLCGMOEIDMSDWKSTIY 766
QY 644 KHTCPSNIVKFWKAVFEFFDERRERLLOFVTVGSSRVLPQGFALQGAAGRPFLTFHQI 703
DB 767 RHYTKNSKQIOWVQVQVWENDNEKRILOFVTVGTCRLPVGGFAELIGSGNQKFCIDKV 826
QY 704 DACTNNLPKATCFNRIDIPYESYKLYKLLTAEETCGFAVE 748
DB 827 GKET-WLPSRHTCFNRDLDPYKSYEQLEKLYAIEETEGFQE 870
RESULT 13
ID_WP1_HUMAN STANDARD; PRT; 922 AA.
AC Q9HOM0; O00307; Q96BP4;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Nedd4-like ubiquitin-protein ligase WBP1 (EC 6.3.2.-) (WW domain-
containing protein 1) (Atropin-1 interacting protein 5) (AIP5).
GN WBP1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RC TISSUE=Testis;
RX MEDLINE=21154917; PubMed=11230166;
RA Wiemann S., Weil B., Wellenreuther R., Gassenhuber J., Glassl S.,
RA Ansoerg W., Boecker M., Bloeker H., Bauersachs S., Blum H.,
RA Lauber J., Duesterhoeft A., Beyer A., Koehler K., Strack N.,
RA Mewes H.-W., Ottenwaeider B., Obermaier B., Tampe J., Heubner D.,
RA Wambutt R., Korn B., Klein M., Foutska A.;
RA "Towards a catalog of human genes and proteins: sequencing and
RT analysis of 500 novel complete protein coding human cDNAs.";
RL Genome Res. 11:422-435(2001).
RN [2]
RP SEQUENCE FROM N.A. (ISOFORMS 1; 2; 3; 4; 5 AND 6), AND TISSUE
RP SPECIFICITY.
RC TISSUE=Breast cancer;
RX MEDLINE=21638010; PubMed=11779188;
RA Flaszka M., Gorman P., Roylance R., Canfield A.E., Baron M.;
RT "Alternative splicing determines the domain structure of WBP1, a
RT Nedd4 family protein.";
RL Biochem. Biophys. Res. Commun. 290:431-437(2002).
RN [3]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RC TISSUE=Lung, and Testis;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Scapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Udwin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellaro N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fanev J., Helton A., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
RA Whitting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,

Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smallos D.E.,
Schnerch A., Schein J.B., Jones S.J.N., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [4]
RP SEQUENCE OF 191-870 FROM N.A., AND INTERACTION WITH WBP1; WBP2;
RP SCNN1A; SCNN1B AND SCNN1G.
RC TISSUE=Bone marrow, and Brain;
RX MEDLINE=97313427; PubMed=9169421;
RA Pirozzi G., McConnell S.J., Uveges A.J., Carter J.M., Sparks A.B.,
RA Kay B.K., Fowlkes D.M.;
RT "Identification of novel human WW domain-containing proteins by
RT cloning of ligand targets";
RL J. Biol. Chem. 272:14611-14616(1997).
RN [5]
RP INTERACTION WITH DRPLA, AND TISSUE SPECIFICITY.
RX MEDLINE=98313405; PubMed=9647693;
RA Wood J.D., Yuan J., Margolis R.L., Colomer V., Duan K., Kushi J.,
RA Kaminsky Z., Kleiderlein J.J. Jr., Sharp A.H., Ross C.A.;
RT "Atrophin-1, the DRPLA gene product, interacts with two families of
RT WW domain-containing proteins.";
RL Mol. Cell. Neurosci. 11:149-160(1998).
RN [6]
RP INTERACTION WITH PIII.
RX MEDLINE=22338357; PubMed=12450395;
RA Galanter R., Gout E., Lortat-Jacob H., Wood J., Chroboczek J.;
RT "Adenovirus protein involved in virus internalization recruits
RT ubiquitin-protein ligases.";
RL Biochemistry 41:14299-14305(2002).
RN [7]
RP X-RAY CRYSTALLOGRAPHY (2.2 ANGSTROMS) OF 546-917, MUTAGENESIS OF
RP GLU-614; HIS-621; ASP-675; GLU-798; MET-804; ARG-845; GLN-848
RP AND ARG-855, AND FUNCTION.
RX MEDLINE=22423789; PubMed=12535537;
RA Verdecia M.A., Joazeiro C.A.P., Wells N.J., Ferrer J.-L., Bowman M.E.,
RA Hunter T., Noel J.P.;
RT "Conformational flexibility underlies ubiquitin ligation mediated by
RT the WBP1 HECT domain E3 ligase.";
RL Mol. Cell 11:249-259(2003).
CC -!- FUNCTION: E3 ubiquitin-protein ligase which accepts ubiquitin from
CC an E2 ubiquitin-conjugating enzyme in the form of a thioester and
CC then directly transfers the ubiquitin to targeted substrates.
CC -!- PATHWAY: Ubiquitin conjugation; third step.
CC -!- SUBUNIT: Binds KLF2 (By similarity). Binds SCNN1A, SCNN1B, SCNN1G,
CC WBP1, WBP2, DRPLA and adenovirus type 2 PIII.
CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=6;
CC Comment=Additional isoforms seem to exist;
CC Name=1; Synonyms=A;
CC IsoId=Q9HOM0-1; Sequence=Displayed;
CC Name=2; Synonyms=B;
CC IsoId=Q9HOM0-2; Sequence=VSP_007601, VSP_007603;
CC Name=3; Synonyms=C;
CC IsoId=Q9HOM0-3; Sequence=VSP_007602;
CC Name=4; Synonyms=D;
CC IsoId=Q9HOM0-4; Sequence=Not described;
CC Name=5; Synonyms=E;
CC IsoId=Q9HOM0-5; Sequence=Not described;
CC Name=6; Synonyms=F;
CC IsoId=Q9HOM0-6; Sequence=VSP_007600;
CC -!- TISSUE SPECIFICITY: Detected in heart, placenta, pancreas, kidney,
CC liver, skeletal muscle, bone marrow, fetal brain, and at much
CC lower levels in adult brain and lung. Isoforms 1 and 5 predominate
CC in all tissues tested, except in testis and bone marrow, where
CC isoform 5 is expressed at much higher levels than isoform 1.
CC -!- MISCELLANEOUS: A cysteine residue is required for ubiquitin-
CC thiolester formation.
CC -!- SIMILARITY: Contains 1 C2 domain.
CC -!- SIMILARITY: Contains 4 WW domains.
CC -!- SIMILARITY: Contains 1 HECT-type E3 ubiquitin-protein ligase

domain.

 This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

 DR EMBL; A1136739; CAB56673.1; -;
 DR EMBL; AY043361; AAK94668.1; -;
 DR EMBL; BC015380; AAH15380.1; -;
 DR EMBL; BC036065; AAH36065.1; -;
 DR EMBL; U96113; AAC51324.1; -;
 DR PDB; 1ND7; 23-SEP-03.
 DR MIM; 602307; -;
 DR GO; GO:0000151; C:ubiquitin ligase complex; NAS.
 DR GO; GO:0005515; C:protein binding; IPI.
 DR GO; GO:0004842; F:ubiquitin-protein ligase activity; TAS.
 DR GO; GO:0007417; P:central nervous system development; NAS.
 DR GO; GO:0030324; P:lung development; ISS.
 DR GO; GO:0016481; P:negative regulation of transcription; ISS.
 DR GO; GO:0016567; P:protein ubiquitination; TAS.
 DR GO; GO:0007165; P:signal transduction; NAS.
 DR GO; GO:0030217; P:T-cell differentiation; ISS.
 DR GO; GO:0046718; P:viral entry; TAS.
 DR InterPro; IPR000008; C2.
 DR InterPro; IPR008973; C2_CaLB.
 DR InterPro; IPR000569; HECT_domain.
 DR InterPro; IPR002349; WW.
 DR InterPro; IPR001202; WW_rsp5_WWP.
 DR Pfam; PF00168; C2; 1.
 DR Pfam; PF06632; HECT; 1.
 DR Pfam; PF00397; WW; 4.
 DR PRINTS; PR00403; WWDOMAIN.
 DR SMART; SM00239; C2; 1.
 DR SMART; SM00119; HECTC; 1.
 DR SMART; SM00456; WW; 4.
 DR PROSITE; PS00499; C2_DOMAIN_1; FALSE_NEG.
 DR PROSITE; PS50004; C2_DOMAIN_2; 1.
 DR PROSITE; PS50237; HECT; 1.
 DR PROSITE; PS01159; WW_DOMAIN_1; 4.
 DR PROSITE; PS50020; WW_DOMAIN_2; 2.
 KW Ub1 conjugation pathway; Ligase; Repeat; Alternative splicing;
 3D-structure.
 FT DOMAIN 5 98 C2 DOMAIN.
 FT DOMAIN 349 382 WW 1.
 FT DOMAIN 381 414 WW 2.
 FT DOMAIN 456 489 WW 3.
 FT DOMAIN 496 529 WW 4.
 FT DOMAIN 588 922 HECT.
 FT BINDING 890 890 MISSING (in isoform 6).
 FT VARSPLIC 23 240 /FTID=VSP 007600.
 FT VARSPLIC 112 120 LERVKEOLK -> CWLLKARVE (in isoform 2).
 FT VARSPLIC 112 242 LERVKEOLKSLRNKNGIAQTGELTVLDGLVIEQENTNC
 FT SPSPTIEIQENGDAHENGEPSSARTARLAVEGIDNHV
 FT PFTSLVQSCCYVNGDTPSSQVAARPKNTAPKPLA
 FT SEPADTV -> F (in isoform 3).
 FT /FTID=VSP 007602.
 FT Missing (in isoform 2).
 FT /FTID=VSP 007603.
 FT E->A; REDUCES UBIQUITIN TRANSFER.
 FT H->A; STRONGLY REDUCES UBIQUITIN TRANSFER.
 FT D->A; REDUCES UBIQUITIN TRANSFER.
 FT E->A; REDUCES UBIQUITIN TRANSFER.
 FT STRONGLY REDUCES UBIQUITIN TRANSFER; WHEN
 FT ASSOCIATED WITH A-845.
 FT X->P; STRONGLY REDUCES UBIQUITIN
 FT TRANSFER; WHEN ASSOCIATED WITH P-806.
 FT

FT	MUTAGEN	806	806	E->P; STRONGLY REDUCES UBIQUITIN TRANSFER; WHEN ASSOCIATED WITH P-804. R->A; NO EFFECT. Q->A; ABOLISHES UBIQUITIN TRANSFER; WHEN
FT	MUTAGEN	845	845	
FT	MUTAGEN	848	848	
Query Match 32.1%; Score 1297; DB 1; Length 922; Best Local Similarity 33.3%; Pred. No. 1.2e-84; Matches 317; Conservative 125; Mismatches 268; Indels 242; Gaps 22;				
QY	2	SNFRRNGPVKRLRLTVLCAKMLVKDDFFRLPPFAKVVVDSGQCHSDTVK--NTLDP	59	
DB	8	SDTSNNHSGRLQVTVSSAKLRKKWFGTA-IYTEVVVDG---EITKAKSSSSNP	62	
QY	60	KWQHYDLYIKGSDSVTISWNHK-----K	84	
DB	63	KWDEQLTVNVTPQTLEFQVWSHRTLKADALLGKATIDLKALLIHNRLERVKELKLS	122	
QY	85	IKKQAGFLG-----95		
DB	123	LENKNGIAQTGELTVLDGLVIEQENTNCSSPTIEIQENGDAHENGEPSSARTARLA	182	
QY	96	-----CVLLSNAIN-----RLKDTGYQL-----115		
DB	183	VEGTNGIDNHVPTSLVQSCCYVNGDTPSSQVAARPKNTAPKPLASEPADTV	242	
QY	116	--DLCKLGPNDNTVRGQIVVLSQR-----DRIGTGGQVDC-----151		
DB	243	NGESSFAPTDNASVTGTPVSEENALSPNCTSTTVEDPPVQBIILTSSENNECIPSTSAE	302	
QY	152	-----SRLED-----NDLPDQWEERTAS	170	
DB	303	LESEARSILEPDTSNRSRSSFSAAKSRPDGCMDPVROOSGNANTETLPSCWEQRDPH	362	
QY	171	GRIQYLNHITRTQWER--PTRPASEYSPGRPLSCFVDENTPISGTNGATC-----GQ	222	
DB	363	GRITYVDHNRITTTTWERPQLPGWERRVDRRVVVDHNRITTTTQWRFTMESVRNEQ	422	
QY	223	SSDPRLAERVRQRHNTM--SRTHLHTPPD-----LPEGVEQRTTQGGVYFLHTQGV	277	
DB	423	WQSRNVLQAGMQQFNQVLYSASMLAENDPYGLPPGWEKRVDSRDYVFNHNTKT	482	
QY	278	TWEDPRVPRDLSNINCBELPLPGWEIRNTAGRVYFVDHNNRTTQFDPRLSANLHV	337	
DB	483	QWEDPRT--QGLQN--EE--PLPEGHEIYTRGVYFVDHNRITTTTQKDR-----528		
QY	338	LNQNLKQOQQQVSVLSCLPDTECLTVPRYKRLVOKLILRQELSQOQOAGHCRLEV	397	
DB	529	-NGKSSVTGGPQIA-----YERGFERWKLAFRY-LCQSNALPSHVKNV	571	
QY	398	SREEIFEESYRQVMKRPKDLWKRLMKRGERGLDYGVARSEWLYLSHEMLNPYGLF	457	
DB	572	SRTLFDSPQQIMALKPYDLRRRLVYFRGEBGLDYGGLAREWFFLLSHEVLNPLYCLF	631	
QY	458	QYSRDDIYTLQINPDSAVNPEHLSYFFHVRIMGMAVFHGYIDGGFTLPFYKQLAGSI	517	
DB	632	EYAGKNYCIQINPASTINPDHLSYFCFGRFIAMALFHGKFDITGSLFPFYKMLSKKL	691	
QY	518	TLDMELVDPDLNSLVILENDITGV-LDHIFCVENHAYGEIIQHELKPKGKSIIPNEE	576	
DB	692	TIKDLSEIDTEFFYNLSIWRDNNIEECGLEMYFSDMEILGKVTSHDLKGGSNILVTEE	751	
QY	577	NKXEYVRLVYVWFRGLGIAQFLALQKGFNEVPIQHLKLTFOBEKELELIICGLGKIDVD	636	
DB	752	NKDEYGLMTEWFRSGVQEQTKAFLDGNEVVPLQVLFDEKELEVMCGVQEVLDAD	811	
QY	637	WKYNRLKCHTDPDSNIVKFWKAVBFDEBERRARLLQFVTGSSRVLPQGFALQAGAPR	696	
DB	812	WQNTVYRHYTRNSKQIIFWQFVKETDNEVRMLLQFVTGTCLPLGGFAELMGSGNQ	871	
QY	697	LFTIHQIDACTNNLPKAHTCFNRIDIPPYVESKLYEKLITAEETCGGAVE	748	
DB	872	KFCIEKVGKDT-WLPRSHTCFNRDLDPPIKSYEQLEKELFLAIEETEGGQE	922	


```

Db 454 YTEGVRYFVDNTRTTTTFKPRPGFE-----SGTKQSGPG 489
QY 367 RYKRLVOKLKITRLQELSQOQACHRIEVSREBIFESYQVQWMPKDLWLKLMKF 426
Db 490 AYDRSFRKXHYOFR-FLCHSNALPSHVKSIVSRQTLFEDSFQQIMWKPDLRRRLYIM 548
QY 427 RGEGLDYGAVAREWLYLLSHMLNPPYGLFOYSRDDIYTIQINPDSAVNPEHLSYHFV 486
Db 549 RGEGLDYGAVAREWLYLLSHMLNPPYGLFOYSRDDIYTIQINPDSAVNPEHLSYHFV 608
QY 487 GRIMGMAVPHGYIDGGTLPYKOLLGKSTLDDMELAVDPDLHNSLVWILENDITGV-L 545
Db 609 GFIAWALYHGVKFDITGTFILFYKMLNKRPTLKDLSIDFEFYNSIVWIKENLDECG 668
QY 546 DHTFCVENAYGEIIQHLEKPNKSGIPVNEENKKEYLYVNVNRLRGIEAQFLALQKGF 605
Db 669 ELFFIQDMELLGKVTTHLEKGENIRVTEENKEEYIMLLTDWRTGRCVGEQTKAFLDGF 728
QY 606 NEVIFQHLKTFDEKELELITCGLKIDVNDWKNVTRLKHCTPDPSNIVWFKWAVEFDE 665
Db 729 NEVAPLEWIRYFDEKELEMLCGMQEIDMSDWQKNVIRYHYTKSKQIQFWQVVKEND 788
QY 666 ERARLLQFVGTSSRVPLQGFKALQGAAGPRLFTTHQIDACTNNLPKHAHTCFNRIDIPPY 725
Db 789 EKRIQLQFVGTGTCRLPVGGFAELIGSNGPKFCIDRVGKET-WLPRSHTCFNRIDIPPY 847
QY 726 ESYEKLYPEKLLTAIBETCGFAVE 748
Db 848 KSYEQLEKLLYAIBETEFGQE 870

```

RESULT 15

```

WWP1_MOUSE
ID WWP1_MOUSE STANDARD; PRT; 918 AA.
AC Q8BZ23; Q8BIV9; Q8VDP8;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Nedd-4-like ubiquitin-protein ligase WWP1 (EC 6.3.2.-) (WW domain-
  containing protein 1).
GN WWP1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
XN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Testis;
RX MEDLINE=22354683; PubMed=12466851;
RA Okazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S.,
RA Nikaide I., Oeato N., Saito R., Suzuki H., Yamanaoka I., Kiyosawa H.,
RA Yagi K., Tomaru Y., Hasegawa Y., Nogami A., Schonbach C., Gojebori T.,
RA Balcaralli R., Hill D.P., Bult C., Hume D.A., Quackenbush J.,
RA Schriml L.M., Knäpflin A., Matsuda H., Batalov S., Beisel K.W.,
RA Blake J.A., Bradt D., Bruscia V., Chothia C., Corbani L.E., Cousins S.,
RA Dalia E., Dragani T.A., Fletcher C.F., Forrest A., Frazer K.S.,
RA Gaasterland T., Gariboldi M., Gissi C., Godzik A., Gough J.,
RA Grimmond S., Gustincich S., Hirokawa N., Jackson I.J., Jarvis E.D.,
RA Kanai A., Kawaji H., Kawasawa Y., Kedzierski R.M., King B.L.,
RA Kongaya A., Kurochkin I.V., Lee Y., Lenhard B., Lyons P.A.,
RA Maglott D.R., Maltas L., Marchionni L., McKenzie L., Miki H.,
RA Nagashima T., Nunata K., Okido T., Pavan W.J., Perteu G., Pesole G.,
RA Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramachandran S.,
RA Ravasi T., Reed J.C., Reed D.J., Reid J., Ring B.Z., Ringwald M.,
RA Sandelin A., Schneider C., Semple C.A., Setou M., Shimada K.,
RA Sulelana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita M.,
RA Verardo R., Wagner L., Wahnstedt C., Wang Y., Watanabe Y., Wells C.,
RA Wilming L.G., Wysshaw-Boris A., Yanagisawa M., Yang I., Yang L.,
RA Yuan Z., Zavolan M., Zhu Y., Zimmer A., Carninci P., Hayatsu N.,
RA Hirozane-Kishikawa T., Konno H., Nakamura M., Sakazume N., Sato K.,
RA Shiraki T., Waki K., Kawai J., Aizawa K., Arakawa T., Fukuda S.,
RA Hara A., Hashizume W., Imotani K., Ishii Y., Itoh M., Kagawa I.,

```

```

RA Miyazaki A., Sakai K., Sasaki D., Shibata K., Shinagawa A.,
RA Yasunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J.,
RA Birney E., Hayashizaki Y.;
RT "Analysis of the mouse transcriptome based on functional annotation of
  60,770 full-length cDNAs.";
RL Nature 420:563-573 (2002).
RN
RP SEQUENCE OF 639-918 FROM N.A.
RC STRAIN=FVB/N; TISSUE=Breast cancer, and Colon;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Hsieh F.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Haefl F.,
RA Diatchenko L., Marusina K., Farmer A.F., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Udwin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.A., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
RA Besak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Munz D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Kzywinski M.I., Skalska U., Smalish D.E.,
RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
  human and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
RN
RP INTERACTION WITH KLF2.
RX MEDLINE=21369920; PubMed=11375995;
RA Conkright M.D., Wani M.A., Lingrel J.B.;
RT "Lung Krueppel-like factor contains an autoinhibitory domain that
  regulates its transcriptional activation by binding WWP1, an E3
  ubiquitin ligase.";
RL J. Biol. Chem. 276:29299-29306 (2001).
CC -1- FUNCTION: E3 ubiquitin-protein ligase which accepts ubiquitin from
  an E2 ubiquitin-conjugating enzyme in the form of a thioester and
  then directly transfers the ubiquitin to targeted substrates (By
  similarity).
CC -1- SUBUNIT: Ubiquitin conjugation; third step.
CC -1- SUBUNIT: Binds SCNN1A, SCNN1B, SCNN1G, WBP1, WBP2, DRPLA and
  adenovirus type 2 pIII (By similarity). Binds KLF2.
CC -1- MISCELLANEOUS: A cysteine residue is required for ubiquitin-
  thioester formation (By similarity).
CC -1- SIMILARITY: Contains 1 C2 domain.
CC -1- SIMILARITY: Contains 4 WW domains.
CC -1- SIMILARITY: Contains 1 HECT-type E3 ubiquitin-protein ligase
  domain.
CC -1- CAUTION: Ref.1 (BAC38473) sequence differs from that shown after
  position 391; the second part of that clone maps to another
  chromosome.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
  between the Swiss Institute of Bioinformatics and the EMBL outstation -
  the European Bioinformatics Institute. There are no restrictions on its
  use by non-profit institutions as long as its content is in no way
  modified and this statement is not removed. Usage by and for commercial
  entities requires a license agreement (See http://www.isb-sib.ch/announce/
  or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AK031338; BAC28168.1; -
CC EMBL; AK023346; BAC38473.1; ALT SEQ.
CC EMBL; BC021470; ABA21470.1; ALT_INIT.
CC EMBL; BC051405; ABA51405.1; -
CC EMBL; MGI:1861728; Wppl.
CC MGD; MGI:1861728; Wppl.
CC GO; GO:0000151; C:ubiquitin ligase complex; ISS.
CC GO; GO:0005515; F:protein binding; IPI.
CC GO; GO:0004942; F:ubiquitin-protein ligase activity; ISS.
CC GO; GO:0007417; P:central nervous system development; ISS.
CC GO; GO:0030324; P:lung development; TAS.

```



```

Db      856  GFAELMGSGNGPKFCIEKVGGKDT-WLPRSTCFNRLDPYKSVQLEKLLFAIBETEG 91L
      745  FAVE 748
      |
Db      915  FGQE 918

Search completed: September 21, 2004, 07:44:22
Job time : 23.2889 secs

```

GenCore version 5.1.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 21, 2004, 07:47:37 ; Search time 95.0891 Seconds
(without alignments)
2526.134 Million cell updates/sec

Title: US-10-009-945-4

Perfect score: 4038

Sequence: 1 MSNPGRRNGPVKRLTVLC.....EKLYEKLTAETGCGFAVE 748

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1342398 seqs, 321133274 residues

Total number of hits satisfying chosen parameters: 1342398

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA.*

- 1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pcp.*
- 2: /cgn2_6/ptodata/2/pubpaa/PCRT_NEW_PUB.pcp.*
- 3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pcp.*
- 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pcp.*
- 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pcp.*
- 6: /cgn2_6/ptodata/2/pubpaa/PCRTUS_PUBCOMB.pcp.*
- 7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pcp.*
- 8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pcp.*
- 9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pcp.*
- 10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pcp.*
- 11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pcp.*
- 12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pcp.*
- 13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pcp.*
- 14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pcp.*
- 15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pcp.*
- 16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pcp.*
- 17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pcp.*
- 18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pcp.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	4031	99.8	748	US-10-021-660-81	Sequence 81, Appl
2	4027	99.7	804	US-10-220-120-407	Sequence 407, App
3	3939.5	97.6	735	US-10-313-955-2	Sequence 2, Appli
4	3001	74.3	722	US-10-097-534-14	Sequence 14, Appl
5	2775	68.7	514	US-09-764-875-819	Sequence 819, App
6	1564	38.7	766	US-10-313-955-4	Sequence 4, Appli
7	1530.5	37.9	832	US-10-032-585-7296	Sequence 7296, Ap
8	1453.5	36.0	869	US-10-128-714-8162	Sequence 8162, Ap
9	1370.5	33.9	911	US-10-205-823-279	Sequence 279, App
10	1355.5	33.6	725	US-10-185-050-126	Sequence 126, App
11	1354.5	33.5	854	US-10-205-823-277	Sequence 277, App
12	1354.5	33.5	927	US-10-097-534-15	Sequence 15, Appl
13	1353.5	33.5	995	US-10-097-534-9	Sequence 9, Appli
14	1353.5	33.5	995	US-10-205-823-275	Sequence 275, App
15	1353	33.5	834	US-10-313-955-6	Sequence 6, Appli

16	1340.5	33.2	759	14	US-10-128-714-3162	Sequence 3162, Ap
17	1338.5	33.1	854	14	US-10-287-218-3	Sequence 3, Appli
18	1338.5	33.1	854	16	US-10-474-291-3	Sequence 3, Appli
19	1301	32.2	739	12	US-10-182-936A-89	Sequence 89, Appl
20	1301	32.2	739	14	US-10-097-534-10	Sequence 10, Appl
21	1301	32.2	739	15	US-10-374-979-89	Sequence 89, Appl
22	1301	32.2	752	10	US-09-919-039-235	Sequence 235, App
23	1298.5	32.2	870	14	US-10-097-534-12	Sequence 12, Appl
24	1298.5	32.2	906	14	US-10-185-050-48	Sequence 48, Appl
25	1228	30.4	898	12	US-10-188-186-114	Sequence 114, App
26	1209	29.9	474	10	US-09-774-639-371	Sequence 371, App
27	1209	29.9	474	10	US-09-969-730-249	Sequence 249, App
28	1209	29.9	474	16	US-10-621-363-249	Sequence 249, App
29	1134.5	28.1	683	14	US-10-185-050-46	Sequence 46, Appl
30	1134.5	28.1	684	14	US-10-097-534-11	Sequence 11, Appl
31	1050	26.0	1562	14	US-10-097-534-13	Sequence 13, Appl
32	1019.5	25.2	380	14	US-10-307-956-32	Sequence 32, Appl
33	1008.5	25.0	375	14	US-10-307-956-31	Sequence 31, Appl
34	869	21.5	733	14	US-10-097-534-16	Sequence 16, Appl
35	835.5	20.7	1094	14	US-10-043-487-300	Sequence 300, App
36	835.5	20.7	1488	14	US-10-043-487-285	Sequence 285, App
37	835.5	20.7	2011	16	US-10-408-765A-767	Sequence 767, App
38	832.5	20.6	1104	12	US-10-424-599-146002	Sequence 146002,
39	828.5	20.5	1843	12	US-10-424-599-146004	Sequence 146004,
40	828.5	20.5	3647	16	US-10-437-963-119793	Sequence 119793,
41	791	19.6	157	12	US-10-211-462-107	Sequence 107, App
42	775.5	19.2	277	9	US-09-925-300-1628	Sequence 1628, Ap
43	771	19.1	973	16	US-10-437-963-178922	Sequence 178922
44	733	18.2	358	14	US-10-268-036-5	Sequence 5, Appli
45	733	18.2	358	16	US-10-391-364-33	Sequence 33, Appl

ALIGNMENTS

RESULT 1

US-10-021-660-81
; Sequence 81, Application US/10021660
; Publication No. US20030152926A1
; GENERAL INFORMATION:
; APPLICANT: Murray, Richard
; APPLICANT: Glynn, Richard
; APPLICANT: Watson, Susan R.
; APPLICANT: EOS Biotechnology, Inc.
; TITLE OF INVENTION: No. US20030152926A1 Methods of Diagnosis of Angiogenesis,
; TITLE OF INVENTION: Compositions and Methods of Screening for Angiogenesis
; TITLE OF INVENTION: Modulators
; FILE REFERENCE: 018501-000710US
; CURRENT APPLICATION NUMBER: US/10/021,660
; CURRENT FILING DATE: 2001-12-06
; PRIOR APPLICATION NUMBER: US/09/784,356
; PRIOR FILING DATE: 2001-02-14
; PRIOR APPLICATION NUMBER: US 09/637,977
; PRIOR FILING DATE: 2000-08-11
; NUMBER OF SEQ ID NOS: 135
; SOFTWARE: Fast-Seq for Windows Version 3.0
; SEQ ID NO 81
; LENGTH: 748
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-021-660-81

Query Match 99.8%; Score 4031; DB 14; Length 748;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 747; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MSNPGRRNGPVKRLTVLCACNLVKKDFRLPDPFAKVVDGSGGCHSDTIVKNTLDPK 60

Db 1 MSNPGRRNGPVKRLTVLCACNLVKKDFRLPDPFAKVVDGSGGCHSDTIVKNTLDPK 60

Qy 61 WNGHYDLYIGKSDSVTISVWNHKKIKHKKOGAGFLGCVRLLSNAINRLKDTGYORLDCKL 120

Db 61 WNGHYDLYIGKSDSVTISVWNHKKIKHKKOGAGFLGCVRLLSNAINRLKDTGYORLDCKL 120

121 GPNDNTRVQGVVSLQSRDRIGTGGVVDVCSLFDNDLPDGWEERTASGRIOYLNHIT 180
Db 121 GPNDNTRVQGVVSLQSRDRIGTGGVVDVCSLFDNDLPDGWEERTASGRIOYLNHIT 180
181 RTTQWERPTRPASEYSSPGRLSCFVDENTPISGTNGATCGOSSDPRLAERRVRSQRHN 240
Db 181 RTTQWERPTRPASEYSSPGRLSCFVDENTPISGTNGATCGOSSDPRLAERRVRSQRHN 240
241 YMSRTHLHTPPDLPEGVEORTTQGGVYFLHTGTGVSTWHDPRVPRDLNINCEELGFLP 300
Db 241 YMSRTHLHTPPDLPEGVEORTTQGGVYFLHTGTGVSTWHDPRVPRDLNINCEELGFLP 300
301 PGWEIRNTATGRVYFVDHNNRTTQTDPRLSANLHLVLRNQLKQDQOQQVVSCLPDDT 360
Db 301 PGWEIRNTATGRVYFVDHNNRTTQTDPRLSANLHLVLRNQLKQDQOQQVVSCLPDDT 360
361 ECLTVPRYKRDVLQVKILRQELSQOQPOAGHCRIEVSREEIFEESYRQVMKRPKDLWK 420
Db 361 ECLTVPRYKRDVLQVKILRQELSQOQPOAGHCRIEVSREEIFEESYRQVMKRPKDLWK 420
421 RLMTKFRGEGLDYGVARWELVLLSHEMLNPPYGLFOYSRDDIYTLQINPDSAVNPEHL 480
Db 421 RLMTKFRGEGLDYGVARWELVLLSHEMLNPPYGLFOYSRDDIYTLQINPDSAVNPEHL 480
481 SYPHFVGRIMGMAVPHGYIDGGFTLPFYKOLLGKSTITLDDMELVDPDLHNSLVWILEND 540
Db 481 SYPHFVGRIMGMAVPHGYIDGGFTLPFYKOLLGKSTITLDDMELVDPDLHNSLVWILEND 540
541 ITGVLDHTFCVEHNAYPEIIOHELKPKNGKSTPVNEENKKEYVRLVYNNWRFLRGIEAQFLA 600
Db 541 ITGVLDHTFCVEHNAYPEIIOHELKPKNGKSTPVNEENKKEYVRLVYNNWRFLRGIEAQFLA 600
601 LQKGFNEVPOHLKTFDEKELELIICGLGKIDVNDKVNTRLKHCCTPDSNIVKWFKAV 660
Db 601 LQKGFNEVPOHLKTFDEKELELIICGLGKIDVNDKVNTRLKHCCTPDSNIVKWFKAV 660
661 EFFDEERRARLQFVTGSSRVPLOGFKALOGAAGPRFTTHQIDACTNNLPKHAHTCFNRI 720
Db 661 EFFDEERRARLQFVTGSSRVPLOGFKALOGAAGPRFTTHQIDACTNNLPKHAHTCFNRI 720
721 DIPVYESYKLYEKLTAIESTCGFAVE 748
Db 721 DIPVYESYKLYEKLTAIESTCGFAVE 748
RESULT 2
US-10-220-120-407
; Sequence 407, Application US/10220120
; Publication No. US20040048253A1
; GENERAL INFORMATION:
; APPLICANT: INCYTE GENOMICS, INC.
; APPLICANT: PANZER, Scott R.
; APPLICANT: SPIRO, Peter A.
; APPLICANT: BANVILLE, Steven C.
; APPLICANT: SHAH, Purvi
; APPLICANT: CHALUP, Michael S.
; APPLICANT: CHANG, Simon C.
; APPLICANT: CHEN, Alice
; APPLICANT: D'SA, Steven A.
; APPLICANT: AMSHEY, Stefan
; APPLICANT: DAHL, Christopher R.
; APPLICANT: DAM, Tam C.
; APPLICANT: DANIELS, Susan E.
; APPLICANT: DUFOUR, Gerard E.
; APPLICANT: FLORES, Vincent
; APPLICANT: FONG, Willy T.
; APPLICANT: GREENAWALT, Lila B.
; APPLICANT: HILLMAN, Jennifer L.
; APPLICANT: JONES, Anissa L.
; APPLICANT: LIU, Tommy F.
; APPLICANT: ROSEBERRY, Ann M.
; APPLICANT: ROSEN, Bruce H.

APPLICANT: RUSSO, Frank D.
APPLICANT: STOCKREHER, Theresa K.
APPLICANT: DAFFO, Abel
APPLICANT: WRIGHT, Rachel J.
APPLICANT: YAP, Pierre E.
APPLICANT: YU, Jimmy Y.
APPLICANT: BRADLEY, Diana L.
APPLICANT: BRATCHEY, Shawn R.
APPLICANT: CHEN, Wensheng
APPLICANT: COHEN, Howard J.
APPLICANT: HODGSON, David M.
APPLICANT: LINCOLN, Stephen E.
APPLICANT: JACKSON, Stuart
; TITLE OF INVENTION: MOLECULES FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: PT-1113 FCT
; CURRENT APPLICATION NUMBER: US/10/220,120
; CURRENT FILING DATE: 2002-08-26
; PRIOR APPLICATION NUMBER: 60/184,777; 60/184,777; 60/184,777; 60/184,698; 60/184,770; 60/184,774;
60/184,693; 60/184,771; 60/184,813; 60/184,773; 60/184,776;
60/184,769; 60/184,768; 60/184,837; 60/184,697; 60/184,841;
60/184,772; 60/185,213; 60/185,216; 60/204,863; 60/205,221;
60/204,815; 60/203,785; 60/204,821; 60/204,908; 60/204,226;
60/204,525; 60/205,285; 60/205,232; 60/205,323; 60/205,287;
60/205,324; 60/205,286
; PRIOR FILING DATE: 2000-02-24; 2000-02-24; 2000-02-24; 2000-02-24; 2000-02-24;
2000-02-24; 2000-02-24; 2000-02-24; 2000-02-24; 2000-02-24;
2000-02-24; 2000-02-24; 2000-02-24; 2000-02-24; 2000-02-24;
2000-05-17; 2000-05-12; 2000-05-16; 2000-05-16; 2000-05-15;
2000-05-16; 2000-05-17; 2000-05-16; 2000-05-17; 2000-05-17;
2000-05-17; 2000-05-17
; NUMBER OF SEQ ID NOS: 422
; SOFTWARE: PERL Program
; SEQ ID NO 407
; LENGTH: 804
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. US20040048253A1 LG:132147.3.crf3:2000FEB18
US-10-220-120-407
Query Match 99.7%; Score 4027; DB 12; Length 804;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 746; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 MSNPGRRNGPVKRLTLVLCANLVKKDFRLPDPFAKVVVDGSGQCHSTDTVKNITLDPK 60
Db 57 MSNPGRRNGPVKRLTLVLCANLVKKDFRLPDPFAKVVVDGSGQCHSTDTVKNITLDPK 116
QY 61 MNQHYDLYIGKSDSVTISVNNHKKIHKQAGAGFLGCVRLLSNAINRLKDTGYQRLDLCKL 120
Db 117 MNQHYDLYIGKSDSVTISVNNHKKIHKQAGAGFLGCVRLLSNAINRLKDTGYQRLDLCKL 176
QY 121 GPNDNTRVQGVVSLQSRDRIGTGGVVDVCSLFDNDLPDGWEERTASGRIOYLNHIT 180
Db 177 GPNDNTRVQGVVSLQSRDRIGTGGVVDVCSLFDNDLPDGWEERTASGRIOYLNHIT 236
QY 181 RTTQWERPTRPASEYSSPGRLSCFVDENTPISGTNGATCGOSSDPRLAERRVRSQRHN 240
Db 237 RTTQWERPTRPASEYSSPGRLSCFVDENTPISGTNGATCGOSSDPRLAERRVRSQRHN 296
QY 241 YMSRTHLHTPPDLPEGVEORTTQGGVYFLHTGTGVSTWHDPRVPRDLNINCEELGFLP 300
Db 297 YMSRTHLHTPPDLPEGVEORTTQGGVYFLHTGTGVSTWHDPRVPRDLNINCEELGFLP 356
QY 301 PGWEIRNTATGRVYFVDHNNRTTQTDPRLSANLHLVLRNQLKQDQOQQVVSCLPDDT 360
Db 357 PGWEIRNTATGRVYFVDHNNRTTQTDPRLSANLHLVLRNQLKQDQOQQVVSCLPDDT 416
QY 361 ECLTVPRYKRDVLQVKILRQELSQOQPOAGHCRIEVSREEIFEESYRQVMKRPKDLWK 420
Db 417 ECLTVPRYKRDVLQVKILRQELSQOQPOAGHCRIEVSREEIFEESYRQVMKRPKDLWK 476

QY 421 RLMIKRGEGLDYGGVAREWLYLLSHEMLNPPYGLFQYSRDDIYTLQINPDSAVNPEHL 480
DB 477 RLMIKRGEGLDYGGVAREWLYLLSHEMLNPPYGLFQYSRDDIYTLQINPDSAVNPEHL 536
QY 481 SYFHFVGRIMGMAVFHGHYIDGGFTLPFYKQLLGKSIITDDMELVDPDHLNSLVWILEND 540
DB 537 SYFHFVGRIMGMAVFHGHYIDGGFTLPFYKQLLGKSIITDDMELVDPDHLNSLVWILEND 596
QY 541 ITGVLDHTTCVEHNAYGEIIQHELKPNKSIIPVNEENKKEYVRLYVNNRFLRGIEBAQFLA 600
DB 597 ITGVLDHTTCVEHNAYGEIIQHELKPNKSIIPVNEENKKEYVRLYVNNRFLRGIEBAQFLA 656
QY 601 LQKGFNEVPIQHLKTFDSEKELELIICGLGKIDVNDKVNTRLKCHCTPDSNVKWFVKAV 660
DB 657 LQKGFNEVPIQHLKTFDSEKELELIICGLGKIDVNDKVNTRLKCHCTPDSNVKWFVKAV 716
QY 661 EFFDEERRARLLQFVTGSSSRVPLQGFKALQGAAGPRLFTIHOIDACTNNLPKHAHTCFNRI 720
DB 717 EFFDEERRARLLQFVTGSSSRVPLQGFKALQGAAGPRLFTIHOIDACTNNLPKHAHTCFNRI 776
QY 721 DIPPYESYEKLYEKLTAIEETCGFAVE 748
DB 777 DIPPYESYEKLYEKLTAIEETCGFAVE 804

RESULT 3

US-10-313-955-2
; Sequence 2, Application US/10313955
; Publication No. US20030199036A1
; GENERAL INFORMATION:
; APPLICANT: Beach, David H.
; Nefsky, Bradley
; TITLE OF INVENTION: Ubiquitin Ligases, and Uses Related Thereto
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FOLEY, HOAG & ELIOT LLP
; STREET: One Post Office Square
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109-2170

COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA: US/10/313,955
; FILING DATE: 05-Dec-2002
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/09/392,163
; FILING DATE: <Unknown>
; APPLICATION NUMBER: US 08/539,205
; FILING DATE: <Unknown>

ATTORNEY/AGENT INFORMATION:
; NAME: Vincent, Matthew P.
; REGISTRATION NUMBER: 36,709
; REFERENCE/DOCKET NUMBER: CSV-005.01
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 832-1000
; TELEFAX: (617) 832-7000

INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 735 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:

US-10-313-955-2
Query Match 97.6%; Score 3939.5; DB 14; Length 735;

Best Local Similarity 98.0%; Pred. No. 0;
Matches 733; Conservative 0; Mismatches 2; Indels 13; Gaps 1;
QY 1 MSNPGRRNGPVKRLRLTVLCAKNLVKDDFFRLPPFAKVVVDSGQCHSTDTVXNTLDPK 60
DB 1 MSNPGRRNGPVKRLRLTVLCAKNLVKDDFFRLPPFAKVVVDSGQCHSTDTVXNTLDPK 47
QY 61 WQHYDLYIGKSDSVTISVMNHKKIHKKGAGFLGCVRLLSNAINRLKDTGYQRLDLCKL 120
DB 48 WQHYDLYIGKSDSVTISVMNHKKIHKKGAGFLGCVRLLSNAINRLKDTGYQRLDLCKL 107
QY 121 GPNDNDVIRGOIVVLSQSRDRIGTGGQVDCSRLFDNDLPGWEERTASGRTOYLNHIT 180
DB 108 GPNDNDVIRGOIVVLSQSRDRIGTGGQVDCSRLFDNDLPGWEERTASGRTOYLNHIT 167
QY 181 RTTQWERPTTPASYSYSPGRLSCFVDENTPIISGTNGATCGQSDPPLAERRVRSQRHN 240
DB 168 RTTQWERPTTPASYSYSPGRLSCFVDENTPIISGTNGATCGQSDPPLAERRVRSQRHN 227
QY 241 YMSRTHLHTPPDLPEGYEORTTQGGQYVFLHTOTGVSTWHDPRVPRDLSNINCEELGPLP 300
DB 228 YMSRTHLHTPPDLPEGYEORTTQGGQYVFLHTOTGVSTWHDPRVPRDLSNINCEELGPLP 287
QY 301 PGWEIRNTATGRVYFVDHNNRTTQFDPRLSANLHLVNEQNLKDOQQOQVVSCLCPDDT 360
DB 288 PGWEIRNTATGRVYFVDHNNRTTQFDPRLSANLHLVNEQNLKDOQQOQVVSCLCPDDT 347
QY 361 ECLTVPRYKDLVQKLKILRQELSQOQOQAGHCRIEVSREEIFEESYRQVMKRPDLWK 420
DB 348 ECLTVPRYKDLVQKLKILRQELSQOQOQAGHCRIEVSREEIFEESYRQVMKRPDLWK 407
QY 421 RLMIKRGEGLDYGGVAREWLYLLSHEMLNPPYGLFQYSRDDIYTLQINPDSAVNPEHL 480
DB 408 RLMIKRGEGLDYGGVAREWLYLLSHEMLNPPYGLFQYSRDDIYTLQINPDSAVNPEHL 467
QY 481 SYFHFVGRIMGMAVFHGHYIDGGFTLPFYKQLLGKSIITDDMELVDPDHLNSLVWILEND 540
DB 468 SYFHFVGRIMGMAVFHGHYIDGGFTLPFYKQLLGKSIITDDMELVDPDHLNSLVWILEND 527
QY 541 ITGVLDHTTCVEHNAYGEIIQHELKPNKSIIPVNEENKKEYVRLYVNNRFLRGIEBAQFLA 600
DB 528 ITGVLDHTTCVEHNAYGEIIQHELKPNKSIIPVNEENKKEYVRLYVNNRFLRGIEBAQFLA 587
QY 601 LQKGFNEVPIQHLKTFDSEKELELIICGLGKIDVNDKVNTRLKCHCTPDSNVKWFVKAV 660
DB 588 LQKGFNEVPIQHLKTFDSEKELELIICGLGKIDVNDKVNTRLKCHCTPDSNVKWFVKAV 647
QY 661 EFFDEERRARLLQFVTGSSSRVPLQGFKALQGAAGPRLFTIHOIDACTNNLPKHAHTCFNRI 720
DB 648 EFFDEERRARLLQFVTGSSSRVPLQGFKALQGAAGPRLFTIHOIDACTNNLPKHAHTCFNRI 707
QY 721 DIPPYESYEKLYEKLTAIEETCGFAVE 748
DB 708 DIPPYESYEKLYEKLTAIEETCGFAVE 735

RESULT 4
US-10-097-534-14
; Sequence 14, Application US/10097534
; Publication No. US20030049607A1
; GENERAL INFORMATION:
; APPLICANT: GREENER, TSVIKA
; APPLICANT: MOSKOWITZ, HAIM
; APPLICANT: REISS, YUVAL
; APPLICANT: ALROY, IRIS
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE MODULATION OF VIRAL
; FILE REFERENCE: PIV-001.01
; CURRENT APPLICATION NUMBER: US/10/097,534
; CURRENT FILING DATE: 2002-03-12
; PRIOR APPLICATION NUMBER: 60/275,224
; PRIOR FILING DATE: 2001-03-12
; PRIOR APPLICATION NUMBER: 60/308,958

APPLICANT: Rosen et al.
TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
FILE REFERENCE: RJ202
CURRENT APPLICATION NUMBER: US/09/764,875
CURRENT FILING DATE: 2001-01-17
Prior application data removed - consult PALM or file wrapper
NUMBER OF SEQ ID NOS: 1249
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 819
LENGTH: 514
TYPE: PRT
ORGANISM: Homo sapiens
US-09-764-875-819

Query Match 68.7%; Score 2775; DB 11; Length 514;
Best Local Similarity 99.8%; Pred. No. 2.3e-248;
Matches 513; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 235 SORHNTMSRTHLHTPPDLPEGYEQRTTQGGVYFLHTQGVSTWHDPRVPRDLNINCE 294
DB 1 SRRHNTMSRTHLHTPPDLPEGYEQRTTQGGVYFLHTQGVSTWHDPRVPRDLNINCE 60
QY 295 ELGPIPGWEIRNTATGRVYFVDHNNRTTQTPDRLSANLHLVLRNQNLKDDQQQQVVS 354
DB 61 ELGPIPGWEIRNTATGRVYFVDHNNRTTQTPDRLSANLHLVLRNQNLKDDQQQQVVS 120
QY 355 LCPDTECLTVPRYKRDVLQKILROELSQQQPQAGHCRIEVSREIFEEYSRQVMKMR 414
DB 121 LCPDTECLTVPRYKRDVLQKILROELSQQQPQAGHCRIEVSREIFEEYSRQVMKMR 180
QY 415 PKDLWKRLMKFRGEEGLDYGAVAREWLYLLSHEMLNPYYGLFOYSRDDIYTLQINPDSA 474
DB 181 PKDLWKRLMKFRGEEGLDYGAVAREWLYLLSHEMLNPYYGLFOYSRDDIYTLQINPDSA 240
QY 475 VNPEHLSYFHFVGRIMGMAVPHGHYIDGGFTLPYKQLGKSLTLDDELVDLHNSLV 534
DB 241 VNPEHLSYFHFVGRIMGMAVPHGHYIDGGFTLPYKQLGKSLTLDDELVDLHNSLV 300
QY 535 WILENDITGVLDHFCVEHNAAYGIIQHELKPNKSIIPVNEENKKEYVRLYVNWRFMRGI 594
DB 301 WILENDITGVLDHFCVEHNAAYGIIQHELKPNKSIIPVNEENKKEYVRLYVNWRFMRGI 360
QY 595 EAQFLALQKGFNEVLPQHLKTFDEKELELIICGLKIDVNDKVNTRLKHCTPDSNIVK 654
DB 361 EAQFLALQKGFNEVLPQHLKTFDEKELELIICGLKIDVNDKVNTRLKHCTPDSNIVK 420
QY 655 WFKAVEFFDEERRARLLQFVTSRVPLOQFKALQAGAPRLFTIHOIDACTNNLPKAH 714
DB 421 WFKAVEFFDEERRARLLQFVTSRVPLOQFKALQAGAPRLFTIHOIDACTNNLPKAH 480
QY 715 TCFNRIDIPPPESYEKLEKLLTAIEETCGFAVE 748
DB 481 TCFNRIDIPPPESYEKLEKLLTAIEETCGFAVE 514

RESULT 6
US-10-313-955-4
Sequence 4, Application US/10313955
Publication No. US20030199036A1
GENERAL INFORMATION:
APPLICANT: Beach, David H.
Nefsky, Bradley
Caligiuri, Maureen
TITLE OF INVENTION: Ubiquitin Ligases, and Uses Related Thereto
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: FOLEY, HOAG & ELIOT LLP
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109-2170
COMPUTER READABLE FORM:

PRIOR FILING DATE: 2001-07-31
PRIOR APPLICATION NUMBER: 60/340,170
PRIOR FILING DATE: 2001-12-07
NUMBER OF SEQ ID NOS: 71
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 14
LENGTH: 722
TYPE: PRT
ORGANISM: Homo sapiens
US-10-097-534-14

Query Match 74.3%; Score 3001; DB 14; Length 722;
Best Local Similarity 74.6%; Pred. No. 4.1e-269;
Matches 565; Conservative 63; Mismatches 71; Indels 58; Gaps 8;

QY 12 VKLRLTVLCAKLVKDDFFELPDPFAKVVVDGSGQCHSTDTVKNTLDPKWNQHYLDYQK 71
DB 4 IKRLTVLCAKLVKDDFFELPDPFAKVVVDGSGQCHSTDTVKNTLDPKWNQHYLDYQK 63
QY 72 SDSVTSVWNHKKIKKQAGFGCVRLLSNAINRLKDTGYQRLDCKLGPNDNDTVRQ 131
DB 64 TDSVTSVWNHKKIKKQAGFGCVRLLSNAISRLKDTGYQRLDCKLNPSTDAVRQ 123
QY 132 IVVSLQSRDRIGTGGQVVDGSRFLPNDLPDQWEERETASGRIOYLNHITRTTQWERPTR 191
DB 124 IVVSLQSRDRIGTGGVVDGCRGLENE-----GTVY----- 154
QY 192 ASEVSSGRPLSCFVDENTPISGTNGATCG-----QSSPRLAERVRSSQRHNTM 242
DB 155 --EDSGRPLSCFMEPAPYDTSAGAAAGGNCRFVESPQORLOAQLRNPVVRGSL 212
QY 243 ----SRTHLTPDLPGEYEQRTTQGGVYFLHTQGVSTWHDPRVPRDLNINCEELGP 298
DB 213 QTPQNRHGHQSPLPGEYEQRTTQGGVYFLHTQGVSTWHDPRVPRDLNINCEDELGP 272
QY 299 LPPGWEIRNTATGRVYFVDHNNRTTQTPDRLSANLHLVLRNQNLKDDQQQQVVS---S 354
DB 273 LPPGWEIRNTATGRVYFVDHNNRTTQTPDRLSANLHLVLRNQNLKDDQQQQVVS---S 328
QY 355 LCPDTECLTVPRYKRDVLQKILROELSQQQPQAGHCRIEVSREIFEEYSRQVMKMR 414
DB 329 L---EDBELPQVRVERDLQKILROELSLQQQPQAGHCRIEVSREIFEEYSRQVMKMR 385
QY 415 PKDLWKRLMKFRGEEGLDYGAVAREWLYLLSHEMLNPYYGLFOYSRDDIYTLQINPDSA 474
DB 386 PKDLWKRLMKFRGEEGLDYGAVAREWLYLLSHEMLNPYYGLFOYSRDDIYTLQINPDSA 445
QY 475 VNPEHLSYFHFVGRIMGMAVPHGHYIDGGFTLPYKQLGKSLTLDDELVDLHNSLV 534
DB 446 INPDHLSYFHFVGRIMGMAVPHGHYIDGGFTLPYKQLGKSLTLDDELVDLHNSLV 505
QY 535 WILENDITGVLDHFCVEHNAAYGIIQHELKPNKSIIPVNEENKKEYVRLYVNWRFMRGI 594
DB 506 WILENDITGVLDHFCVEHNAAYGIIQHELKPNKSIIPVNEENKKEYVRLYVNWRFMRGI 565
QY 595 EAQFLALQKGFNEVLPQHLKTFDEKELELIICGLKIDVNDKVNTRLKHCTPDSNIVK 654
DB 566 EAQFLALQKGFNEVLPQHLKTFDEKELELIICGLKIDVNDKVNTRLKHCTPDSNIVK 625
QY 655 WFKAVEFFDEERRARLLQFVTSRVPLOQFKALQAGAPRLFTIHOIDACTNNLP 711
DB 626 WFKAVEFFDEERRARLLQFVTSRVPLOQFKALQAGAPRLFTIHOIDACTNNLP 685
QY 712 KAHTCFNRIDIPPPESYEKLEKLLTAIEETCGFAVE 748
DB 686 KAHTCFNRIDIPPPESYEKLEKLLTAIEETCGFAVE 722

RESULT 5
US-09-764-875-819
Sequence 819, Application US/09764875
Publication No. US20040018969A1
GENERAL INFORMATION:

```

; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/313.955
; FILING DATE: 05-Dec-2002
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/09/392.163
; FILING DATE: <Unknown>
; APPLICATION NUMBER: US 08/539,205
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Vincent, Matthew P.
; REGISTRATION NUMBER: 36,709
; REFERENCE/DOCKET NUMBER: CSV-005.01
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 832-1000
; TELEFAX: (617) 832-7000
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 766 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 4:
US-10-313-955-4

Query Match      38.7%; Score 1564; DB 14; Length 766;
Best Local Similarity 41.1%; Pred. No. 1.4e-135;
Matches 338; Conservative 121; Mismatches 232; Indels 132; Gaps 18;

QY 1 MSNPGRRRNGPVKRLTLVLCANLVKDFRLDPDPFAKVVDGSGGCHSTDTVKNLDPK 60
DB 1 MSNSAQR-----RIRTVAAADGLYKRVFRPDPFAVLTVDGE-QTHTTTAKTLNPFY 55
QY 61 WNOHYDLYICKDSVTISVWNHKKIKQAGFLGCVRLLSNAINLKDGYQL--DLG 118
DB 56 WNETFEVNTDNTIALQVFDQKF-KKKGQGLGVINRVGDVLDAIGGDEMLTRDK 114
QY 119 KLPNDNDTVRGQIVVSLQ-----RRTASVLT--- 137
DB 115 K--SNENTVVGKIIINLTAAQTSLQVPSAASGARTQTSITNDPQSSKSSVSBNPA 172
QY 138 -----SRDRTGQGVDCSRLEPN-----DLPDGWEERTASGRQYLNH-TRTQ 184
DB 173 SRPAGFTDNDAPASPASEPRTFSFEDQGRUPPGWERDNLGRYYVDHNTRSTT 232
QY 185 WERPT-----RPASYSPGRPLSCFVDENTPIGNTGATCGSSDPRLAERRVRSQRH 238
DB 233 WIRPNLSSVAGAAAEHLSSAS--SANVTEGVQPSSSNAA-----RRTASVLT--- 279
QY 239 RNYMSRTHLTPDLPPEGVEORTQGGQVYFLHTQTGVSTWHDPRVPRDLSNIN----- 292
DB 280 -----SNATTAGSGELPPGWEQRYTEGRPYFVDHNTRITTTWDPRRQYRSYGGFNAT 335
QY 293 -----CEEGLPLPPGWEIRNTATGRVYFVDHNTRITQTDPLSANLHLVLRNQQLKDX 347
DB 336 IQQQPVSLQGLPLSGWEMLTNTARYFVDHNKTKTTWDDPLPSSL-----DQ 384
QY 348 QQQQVSLCPDTECLTVPRYKRDILVQKLIURQSLSQ--QPOAGHCRIEVSRREEFEE 405
DB 385 -----NVFQYKDRFERKLIYF---LSQPALHPLPGQCHIKVRNHFED 425
QY 406 SYRQVMKMRPKDLWKRLMKFRGEGSLDYGGVAREWLVLLSHEMLNPNYGLFOYSRDDIY 465
DB 426 SYAEIMRQSATDLKRLMKFDGEDGLDYGGSURBYFLLSHEMENPFYCLFEYSVDNY 485
QY 466 TLQINPDSAVNPEHLSYFHFVGRINGMAVFHGHIYDGGFTTLPFYKQLLGKSTITLDDMELV 525
DB 486 TLQINPHSGINPEHLYKFGIRVGLAIFHRFVDFAFVVSFYKMLQKVTLODMESM 545
QY 526 DDDLNSLVLWILENDITGVLDHTFCVEHNAVEIILQHELKPKNGKSIIPVNEENKKEYRLY 585

```

```

DB 546 DAERYRSLWILDNDITGVLDLTFVSVDNCFGEVVTIDLPNGRNIEVTEENKRSYDVLV 605
QY 586 VNWRLRGTEAQFLALQKGFNEVPIQHLLKTFDEKELELIICGLGKIDVNDWKVNLRLKH 645
DB 606 TVW-IQKRIEQFNAFHEGFSELIPEQLINVDERELELLIGGISEIDMEDWKKHKDYRS 664
QY 646 CTPDSNIVKFWKAVEFFDEERRARLLQFVTCSSRVPLQSGKALQAGAPRLFTTHQIDA 705
DB 665 YSENDQIIKFWELMDWESNEKSKRLLOFTTGTSRIPVNGFKDLQSGDGPGRKFTIERAGE 724
QY 706 CTNNLPKARTCFNRIDIPPEYSYEKLYEKLJLTAIBETCGFAVE 748
DB 725 -PNKLPKARTCFNRDLDPYTSKKDLHKLSTAVEETIGFQGE 766

RESULT 7
US-10-032-585-7296
; Sequence 7296, Application US/10032585
; Publication No. US20030180953A1
; GENERAL INFORMATION:
; APPLICANT: Terry, Roemer D.
; APPLICANT: Bo, Jiang
; APPLICANT: Charles, Boone
; APPLICANT: Howard, Bussey
; TITLE OF INVENTION: Gene Disruption Methodologies for Drug Target Discovery
; FILE REFERENCE: 10182-005-999
; CURRENT APPLICATION NUMBER: US/10/032.585
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 8000
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 7296
; LENGTH: 832
; TYPE: PRT
; ORGANISM: Candida albicans
US-10-032-585-7296

Query Match      37.9%; Score 1530.5; DB 14; Length 832;
Best Local Similarity 38.4%; Pred. No. 2.1e-132;
Matches 338; Conservative 120; Mismatches 228; Indels 195; Gaps 21;

QY 6 RRRNGPVKRLTLVLCANLVKDFRLDPDPFAKVVDGSGGCHSTDTVKNLDPK 65
DB 9 RSNNTTINVKVVAASLYKRVFRPDPFAVLTVDGS-QTKTITAKTLNPNWNETF 67
QY 66 DLYIGKSDSV-TISVWNHKKIKQAGFLGCVR-----LLSNAINLKDGYQLDLCK 119
DB 68 N-FOAKEDSILVIQVPOKKF-KKKGQGLGVINRVIGDVIDLSNSESSEITR--DLKK 123
QY 120 LQPNNDNTVRGQIVVSLQSDRIGTGQVVD----- 150
DB 124 --SNENLAVSGKIIVI-SHNRNSNGGVTTATTGTGASSNNIATITSGVNNLRIGSA 180
QY 151 -----CSRLEFDN--- 157
DB 181 TTTANSTAQASSDATVANGSGTSSLPPIGQHPETAATPGGAAGAAASQYSSFED 240
QY 158 ---DLPDGWEERTASGRQYLNHITRTTOWERPTRPASE----- 194
DB 241 QYGLPLPGWERRTDNFGRTYYVDHNSRTTTWQRPALHQSETERGQQRSQSEAEARRHRG 300
QY 195 YSPGR---PLSCFVDENTPIGNTGATCGSS-----DPLAERRVRSQRHN 240
DB 301 RTLPGECSVSL-----PTSGNSITSGNVTVNASGANTPVNPAAVSMAAGATTS 352
QY 241 YMERLHTLTPDLPPEGVEORTQGGQVYFLHTQTGVSTWHDPRVPRDLSNIN----- 292
DB 353 GLG-----ELPFGWEQRTTEGRPYFVDHNTRITTTWDPRRQYIRTFGNTIIQQQ 404
QY 293 -CEEGLPLPGWIRNTATGRVYFVDHNTRITQTDPLSANLHLVLRNQQLKQOQQQ 351
DB 405 PVSQGLPFGWEMRLNTARVYFVDHNTRITTTWDDPLPSSL-----DQ----- 449

```


APPLICANT: Zhao, Xumei
APPLICANT: Anderson, Dustin
TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND
TITLE OF INVENTION: METHODS FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
TITLE OF INVENTION: THERAPY OF PROSTATE CANCER
FILE REFERENCE: MRI-044
CURRENT APPLICATION NUMBER: US/10/205,823
CURRENT FILING DATE: 2002-07-25
PRIOR APPLICATION NUMBER: 60/307,982
PRIOR FILING DATE: 2001-07-25
PRIOR APPLICATION NUMBER: 60/314,356
PRIOR FILING DATE: 2001-08-22
PRIOR APPLICATION NUMBER: 60/325,020
PRIOR FILING DATE: 2001-09-25
PRIOR APPLICATION NUMBER: 60/341,746
PRIOR FILING DATE: 2001-12-12
PRIOR APPLICATION NUMBER: 60/362,158
PRIOR FILING DATE: 2002-03-05
NUMBER OF SEQ ID NOS: 455
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 279
LENGTH: 911
TYPE: PRT
ORGANISM: Homo sapiens
US-10-205-823-279

Query Match 33.9%; Score 1370.5; DB 14; Length 911;
Best Local Similarity 36.7%; Pred. No. 1.8e-117;
Matches 324; Conservative 119; Mismatches 234; Indels 207; Gaps 28;

QY 24 LVKDDFFELP---DPEAKV---VVDGSGOCH---STDVTKNTLPKNOHYDLYIGKSD-S 74
DB 69 LPTDELFPGICDPYKLSIYADENRELALVOTIKTLNPKWNEFFRYFNPNSNR 128
QY 75 VTISVNNHKKIHXQAGFLGCVR-----LLSNAINRLKDTGY 112
DB 129 LLFEVFDENLRD---FLQVDVPLSHLPTEDPTMERYTEKFLLRPRSHKSRVKG 185
QY 113 ORLDLCKLGNNDTVRGQIVSVLSQSDRIGTGQGVVDCSRLPDND-----L 159
DB 186 LRLKMYAPKNG-----GQDENSDQDDMEHGEWVD-----SNDASQHQBELPPPL 235
QY 160 PDGWEERTASGR:QYLNHITRTQWERPT-----RPA 192
DB 236 PPGWEKVDNLGRYYVNNHNRITQMRPSLMDVSVSDNNIRQINQEAHRRFRSRH 295
QY 193 SEYSSP-----GRPLSCFVDENTISGTN-----GATCGOSSDPR-LAE---R 232
DB 296 SEDLEPESEGGDVPEPWET-ISEVNIAGDSLGLALPPPPASPGERTSPQELSELSRR 354
QY 233 VR-----SORHNYMSRTHL-----HTPPD----- 252
DB 355 LQITPDSNGEQFSSLIQREPSRLRSCSVTDVAEQGHLPPGAKDGPVRRRAVKDTLSNPQ 414
QY 253 -----LPEGYEORTTQGOVYELHTQTGVSTWHDPRVP-----RD 287
DB 415 SPQSPYNSPKQHKVTQSPFPQWEMRIAPNGPPFDIDNTKTTTWEDRLKFPVHMS 474
QY 288 LSNICEELGELPQWEIRNTATGRVYFVDHNRITQTDPRLSANLHLVNLNQRLKQ 347
DB 475 KTSLNPNLGLPLPGWEERIHLDGRTFYDHSKITQWEDPRL-----QN----- 519
QY 348 QQQQVSVLCPDDTECLTVPR--YKEDLVOKLILRQELSQQQQAQGHCRIEVSRBEIPE 405
DB 520 -----PAITGPAPVYBREFKQYDYFRKKLKKPADIPNRFENKLNHNIFEE 566
QY 406 SYRQVMKMRPKDLWK-RIMIKFRGEGLDYGGVAREWLYLLSHEMLNPNYYGLFQYSRDI 464
DB 567 SYRRIMSVKRPDLVKARLWIEFESEKGLDYGGVAREWFFLLSKEMENFYGLFEYSATDN 626
QY 465 YTLQINPSAV-NPEHLSYFFVGRINGMVAFHGYIDGQFTLPFYKQLLGKSIITLDDME 523
DB 627 YTLQINPNPNSGLEDHLSYFTFGRVAGLAVFHHKLDGGFFIRPFYKQMLGKQITLNDME 686

QY 524 LVDPDLHNSLVWILENDITGVLDHTTCVHNAYGEIIOHELKPKNGKSIIPVNEENKKEYVR 583
DB 687 SVDSYNSKJWLENDPT-ELDLMECIDDEENFGQTYQVDDLKPNSEIMVTNENKREYID 745
QY 584 LYVNRRLRGIEAQFLALQKGFNEVIPHQLLTKTFDEKELELEIICGLGKIDVNDKVNTRL 643
DB 746 LVTCQRFVNRVQKQMNAFLEGTTELLPIDLIKIFDENELLELMCGLDVVDVNDWRQHSY 805
QY 644 K--HCTPDSNIVKFWKAVEFFDEERRARLLOVGTSSRVPLQGFKALOGAAGPLFTTH 701
DB 806 KNGYC-PNHFVQIWFKAVALLMDAERILLOFVGTTSRVPMNGFAELYSGNSGQPLFTIE 864
QY 702 QIDACTNNLPKATCFNRIDIPPEYSEYKLYEKLITAIETCGF 745
DB 865 QWGS-PEKLPRHTCFNRDLDPYETFEDELREKLLMAVENAQGF 907

RESULT 10

US-10-185-050-126
; Sequence 126, Application US/10185050
; Publication No. US20030077577A1
; GENERAL INFORMATION:
; APPLICANT: Pirozzi, Gregorio
; Kay, Brian K.
; Fowlkes, Dana M.

TITLE OF INVENTION: IDENTIFICATION AND ISOLATION OF NOVEL
POLYPEPTIDES HAVING WM DOMAINS AND METHODS OF USING SAME

NUMBER OF SEQUENCES: 233

CORRESPONDENCE ADDRESS:

ADDRESSEE: PENNIE & EDMONDS LLP

STREET: 1155 Avenue of the Americas

CITY: New York

STATE: New York

COUNTRY: USA

ZIP: 10036-2711

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/10/185,050

FILING DATE: 28-Jun-2002

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/08/826,516

FILING DATE: 03-Apr-1997

ATTORNEY/AGENT INFORMATION:

NAME: MISROCK, S. LESLIE

REGISTRATION NUMBER: 18,872

REFERENCE/DOCKET NUMBER: 1101-208-999

TELECOMMUNICATION INFORMATION:

TELEPHONE: (212) 790-9090

TELEFAX: (212) 896-8864/9741

TELEX: 66141 PENNIE

INFORMATION FOR SEQ ID NO: 126:

SEQUENCE CHARACTERISTICS:

LENGTH: 725 amino acids

TYPE: amino acid

STRANDEDNESS: <Unknown>

TOPOLOGY: unknown

MOLECULE TYPE: peptide

SEQUENCE DESCRIPTION: SEQ ID NO: 126:

US-10-185-050-126

Query Match 33.6%; Score 1355.5; DB 14; Length 725;

Best Local Similarity 43.7%; Pred. No. 3.1e-116;

Matches 280; Conservative 98; Mismatches 200; Indels 63; Gaps 13;

QY 137 QSRDICTGQGVDCSRLFDN---DLPGWEERTASGRLOYNLHITRTQWERPRPAS 193

DB 112 RARSSTVTGSEETPPSVAYVHTTGLPGSGWEERKAKGRTYVYNNHNRITTTTRPINQUA 171

```
194 EYSSPG-----RPLSC---FVDENTPISTNGATCGOSSDPRLAERRVRS 235
172 EDGASGATNNHLEIPEQIRRRSLSSPTVLSAPLEGAKDSPVRAVKDTLSNPQSPQ 231
236 QHRNYSRTHLTPDLPDPEGYEORTTQCGVYFLHTGTGVSTWHDPRVP-----RDLN 290
232 PPSYNSPKQHKVTQSFLPPGWEIRAPNGRFFIDHNTKTTTWEDPRKFPFVHMSKTS 291
291 INCEELGPPPGWEIRNTATGRVYFVDHNNRTTQTPRLSANLHLVLNQNOLKDOQQQ 350
292 LNPDLGPPPGWEIRHLDGRTFYIDHNSKITQWEDPRL-----QN----- 333
351 QVVSCLPDTECLTVPR--YKDLVOKLKLRLQELSQQOQAGHCRLEVSREIFESYR 408
334 -----PAITGPAVPSREFKQKYDYFRKKLKPADIPNRFEMKLRHNNIFESYR 383
409 QVMKRPDLWK-RLMIKFRGEGLDYGVAREWYLLSHEMLNPPYGLFOYSRDDIYTL 467
384 RIMSVKRPDLVKARLWIEFESEKGLDYGVAREWFFLLSKEMENPYGLFEYSATDNYTL 443
468 QINPDSAV-NPEHLSYFHFVGRINGMAVFGHYIDGGFTLPRYKOLGKSTLDDMELVD 526
444 QINPNSGLCNEDHLSYFTFGRVAGLAVPHGKLDGFFIRFFYKMLGKQITLNDMESVD 503
527 PDLHNSLVWILENDITGVLDHTFCVEHNAYGEIIQHELKPNKGKSI PVNEENKKEYVRLYV 586
504 SEYNSLKWILENDPT-ELDLMPCIDENFGQTYQVDLKPNGSEIMVTNENKREYIDLVI 562
587 NWRFLRGIEAQLALQKGFNEVIPHLLKTDFEKELELIICGLKIDVNDKVNTRLK-- 644
563 QMRFVNRVQKQMNAPLEGFTLPIDLIKIFDENELMLGCLGDVDVNDWROHSYKNG 622
645 HCTPDSNIVKFWKAVFEFFDEERRARLLQFVTGSSRVPLQFKALQGAAGPRLFTIHQID 704
623 YC-PNHPVIOQWFKAVLLMDAEKRLLQFVTGTSRVPMNGFAELYGSNGPQLFTIEQWG 681
705 ACTNPLPKAHTCFNRIDIPPYESYEKLYEKLKLTAEETCGF 745
682 S-PEKLPKAHTCFNRIDLPPEYTFEDLQELKLMAVENAOQG 721

RESULT 11
US-10-009-945-4.rapb
; Sequence 277, Application US/10205823
; Publication No. US20030108963A1
; GENERAL INFORMATION:
; APPLICANT: Schlegel, Robert
; APPLICANT: Monahan, John E.
; APPLICANT: Endege, Wilson O.
; APPLICANT: Gannavarapu, Manjula
; APPLICANT: Gorbacheva, Bella
; APPLICANT: Hoersch, Sebastian
; APPLICANT: Kamatkar, Shubhangi
; APPLICANT: Monsey, Angela M.
; APPLICANT: Glatt, Karen
; APPLICANT: Zhao, Xumei
; APPLICANT: Anderson, Dustin
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND
; TITLE OF INVENTION: METHODS FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; TITLE OF INVENTION: THERAPY OF PROSTATE CANCER
; FILE REFERENCE: MRI-044
; CURRENT APPLICATION NUMBER: US/10/205,823
; CURRENT FILING DATE: 2002-07-25
; PRIOR APPLICATION NUMBER: 60/307,982
; PRIOR FILING DATE: 2001-07-25
; PRIOR APPLICATION NUMBER: 60/314,356
; PRIOR FILING DATE: 2001-08-22
; PRIOR APPLICATION NUMBER: 60/325,020
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: 60/341,746
; PRIOR FILING DATE: 2001-12-12
; PRIOR APPLICATION NUMBER: 60/362,158
```

```
194 EYSSPG-----RPLSC---FVDENTPISTNGATCGOSSDPRLAERRVRS 235
172 EDGASGATNNHLEIPEQIRRRSLSSPTVLSAPLEGAKDSPVRAVKDTLSNPQSPQ 231
236 QHRNYSRTHLTPDLPDPEGYEORTTQCGVYFLHTGTGVSTWHDPRVP-----RDLN 290
232 PPSYNSPKQHKVTQSFLPPGWEIRAPNGRFFIDHNTKTTTWEDPRKFPFVHMSKTS 291
291 INCEELGPPPGWEIRNTATGRVYFVDHNNRTTQTPRLSANLHLVLNQNOLKDOQQQ 350
292 LNPDLGPPPGWEIRHLDGRTFYIDHNSKITQWEDPRL-----QN----- 462
351 QVVSCLPDTECLTVPR--YKDLVOKLKLRLQELSQQOQAGHCRLEVSREIFESYR 408
463 -----PAITGPAVPSREFKQKYDYFRKKLKPADIPNRFEMKLRHNNIFESYR 512
409 QVMKRPDLWK-RLMIKFRGEGLDYGVAREWYLLSHEMLNPPYGLFOYSRDDIYTL 467
513 RIMSVKRPDLVKARLWIEFESEKGLDYGVAREWFFLLSKEMENPYGLFEYSATDNYTL 572
468 QINPDSAV-NPEHLSYFHFVGRINGMAVFGHYIDGGFTLPRYKOLGKSTLDDMELVD 526
573 QINPNSGLCNEDHLSYFTFGRVAGLAVPHGKLDGFFIRFFYKMLGKQITLNDMESVD 632
527 PDLHNSLVWILENDITGVLDHTFCVEHNAYGEIIQHELKPNKGKSI PVNEENKKEYVRLYV 586
633 SEYNSLKWILENDPT-ELDLMPCIDENFGQTYQVDLKPNGSEIMVTNENKREYIDLVI 691
587 NWRFLRGIEAQLALQKGFNEVIPHLLKTDFEKELELIICGLKIDVNDKVNTRLK-- 644
692 QMRFVNRVQKQMNAPLEGFTLPIDLIKIFDENELMLGCLGDVDVNDWROHSYKNG 751
645 HCTPDSNIVKFWKAVFEFFDEERRARLLQFVTGSSRVPLQFKALQGAAGPRLFTIHQID 704
752 YC-PNHPVIOQWFKAVLLMDAEKRLLQFVTGTSRVPMNGFAELYGSNGPQLFTIEQWG 810
705 ACTNPLPKAHTCFNRIDIPPYESYEKLYEKLKLTAEETCGF 745
811 S-PEKLPKAHTCFNRIDLPPEYTFEDLQELKLMAVENAOQG 850

RESULT 12
US-10-009-945-4.rapb
; Sequence 15, Application US/10097534
; Publication No. US20030049607A1
; GENERAL INFORMATION:
; APPLICANT: GREENER, TSVIKA
; APPLICANT: MOSKOWITZ, HAIM
; APPLICANT: REISS, YUVAL
; APPLICANT: ALROY, IRIS
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE MODULATION OF VIRAL
; TITLE OF INVENTION: MATURATION
; FILE REFERENCE: PLV-001.01
; CURRENT APPLICATION NUMBER: US/10/097,534
; CURRENT FILING DATE: 2002-03-12
; PRIOR APPLICATION NUMBER: 60/275,224
```

```
131 PRIOR FILING DATE: 2002-03-05
132 NUMBER OF SEQ ID NOS: 455
133 SOFTWARE: FastSeq for Windows Version 4.0
134 SEQ ID NO 277
135 LENGTH: 854
136 TYPE: PRT
137 ORGANISM: Homo sapiens
138 US-10-203-823-277
```

```
Query Match 33.5%; Score 1354.5; DB 14; Length 854;
Best Local Similarity 43.7%; Pred. No. 5e-116;
```

```
Matches 280; Conservative 98; Mismatches 200; Indels 63; Gaps 13;
```


QY 552 EHNAYGEIIQHELKPNKSGIPVNEENKKEVYRLVYNNWFLRGIEAQLALQKGFNEVIPO 611
DB 798 DEENFGOTYQVLLKPNKSGSEIMVTNENKGEYDLVQWRFVNRVQKQNAFLGFTLELLPI 857
QY 612 HLLKTFDEKELELEIIICGLGKIDVNDKVNTRLK--HCTPDSNIVKWFKAVEFFDEBERRA 669
DB 858 DLIKIFDENELELLMGLGVDVNDWRQHSIYKNGYC-PNHPVQWFWKAVLLDAEKRI 916
QY 670 RLLQFTGSSRPVLOQFKALQAGAPRLFTIHQIDACTNNLPKATCFNRIDIPPEYSYE 729
DB 917 RLLQFTGTSRVPVPMNGFAELYGNSGQPLFTIEQWGS-PEKLPRAHCTFNRLDLPPEYTFE 975
QY 730 KLYEKLTAIBETCGF 745
DB 976 DLREKLLMAVENAOGF 991
RESULT 14
US-10-205-823-275
; Sequence 275, Application US/10205823
; Publication No. US20030108963A1
; GENERAL INFORMATION:
; APPLICANT: Schlegel, Robert
; APPLICANT: Monahan, John E.
; APPLICANT: Endege, Wilson O.
; APPLICANT: Gannavarapu, Manjula
; APPLICANT: Gorbacheva, Bella
; APPLICANT: Hoersch, Sebastian
; APPLICANT: Kamatkar, Shubhangi
; APPLICANT: Wonsay, Angela M.
; APPLICANT: Glatt, Karen
; APPLICANT: Zhao, Xumei
; APPLICANT: Anderson, Dustin
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND
; TITLE OF INVENTION: METHODS FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; FILE REFERENCE: MRI-044
; CURRENT APPLICATION NUMBER: US/10/205, 823
; CURRENT FILING DATE: 2002-07-25
; PRIOR APPLICATION NUMBER: 60/307, 982
; PRIOR FILING DATE: 2001-07-25
; PRIOR APPLICATION NUMBER: 60/314, 356
; PRIOR FILING DATE: 2001-08-22
; PRIOR APPLICATION NUMBER: 60/325, 020
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: 60/341, 746
; PRIOR FILING DATE: 2001-12-12
; PRIOR APPLICATION NUMBER: 60/362, 158
; PRIOR FILING DATE: 2002-03-05
; NUMBER OF SEQ ID NOS: 455
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 275
; LENGTH: 995
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-205-823-275
Query Match 33.5%; Score 1353.5; DB 14; Length 995;
Best Local Similarity 44.6%; Pred. NO. 7.9e-116;
Matches 275; Conservative 93; Mismatches 188; Indels 60; Gaps 12;
QY 159 LPDGEERTASGRIOYLNHTRTTOWERPTRPASEYSSPO-----RPLS 203
DB 407 LPSGEERKAGRYYYNNHNRITTTWTPMQAEDGASGASATNSNNHLEIQRPRPS 466
QY 204 C---FVDENTPIGNGATCGOSSDPRLAERVRNSQRHNNMSTHLLHTPPDLPPEGYEQR 260
DB 467 LSSPTVLSAPLEGAKDSFVRRAVNDTLNSPQSPSPYNSPKQKHVQTSFLPPGWEMR 526
QY 261 TTQOQGVFLHQTGVSVWHDPVPE-----RDLSNINCEELGPLPPQWEIRNTATGRVYP 315
DB 527 IAPNGRPFDHTKTTTWEDPRKFPVHMSKTSLSLNDLGPLPPGWERIHLDGRTFY 586

QY 316 VDHNRRTQTTDPRLSANLHLVNRQNLKDDQQQQQVSLCPDDTECLTVPR--YKRDLV 373
DB 587 IDHNSKITQWSDPRL-----QN-----PAITGPAVPYSREPK 618
QY 374 QKLIKRLROELSQQOQACHCRIEVSREEIFEESYRQVWKMRRPKDLWK-RLMIKFRGEEGL 432
DB 619 QKYDYFRKLLKPADIPNRPKMLHRNNIFEESYRIMS VKRPDVLKARLWIEFESEKGL 678
QY 433 DYGVGAREWLVLLSHEMLNPPYGLFQYSRDDIYTLQINPDSAV-NPEHLSYFHFVCRING 491
DB 679 DYGVGAREWVFLLSKEMFNPPYGLFEYSATDNTYTLQINPNSGLCNEDHLSYFTFIGRVAG 738
QY 492 MAVFHGHYIDGGFTLPFYKQLLGKSIITLDDMELVDPDLHNSLVNLENDITGVLDHTFCV 551
DB 739 LAVFHGKLLDGGFFIRPFYKMLGKQITLNDMESVDSEYNSLKWILENDPT-ELDLMFCI 797
QY 552 EHNAYGEIIQHELKPNKSGIPVNEENKKEVYRLVYNNWFLRGIEAQLALQKGFNEVIPO 611
DB 798 DEENFGOTYQVLLKPNKSGSEIMVTNENKGEYDLVQWRFVNRVQKQNAFLGFTLELLPI 857
QY 612 HLLKTFDEKELELEIIICGLGKIDVNDKVNTRLK--HCTPDSNIVKWFKAVEFFDEBERRA 669
DB 858 DLIKIFDENELELLMGLGVDVNDWRQHSIYKNGYC-PNHPVQWFWKAVLLDAEKRI 916
QY 670 RLLQFTGSSRPVLOQFKALQAGAPRLFTIHQIDACTNNLPKATCFNRIDIPPEYSYE 729
DB 917 RLLQFTGTSRVPVPMNGFAELYGNSGQPLFTIEQWGS-PEKLPRAHCTFNRLDLPPEYTFE 975
QY 730 KLYEKLTAIBETCGF 745
DB 976 DLREKLLMAVENAOGF 991
RESULT 15
US-10-313-955-6
; Sequence 6, Application US/10313955
; Publication No. US20030199036A1
; GENERAL INFORMATION:
; APPLICANT: Beach, David H.
; APPLICANT: Caligiuri, Maureen
; APPLICANT: Nefsky, Bradley
; TITLE OF INVENTION: Ubiquitin Ligases, and Uses Related Thereto
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FOLEY, HOAG & ELIOT LLP
; STREET: One Post Office Square
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109-2170
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/313,955
; FILING DATE: 05-Dec-2002
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/09/392,163
; FILING DATE: <Unknown>
; APPLICATION NUMBER: US 08/539,205
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Vincent, Matthew P.
; REGISTRATION NUMBER: 36,709
; REFERENCE/DOCKET NUMBER: CSV-005.01
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 832-1000
; TELEFAX: (617) 832-7000
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 834 amino acids

```
;
;      TYPE: amino acid
;      TOPOLOGY: linear
;      MOLECULE TYPE: protein
;      SEQUENCE DESCRIPTION: SEQ ID NO: 6:
US-10-313-955-6

Query Match      33.5%; Score 1353; DB 14; Length 834;
Best Local Similarity 44.6%; Pred. No. 6.6e-116;
Matches 275; Conservative 93; Mismatches 188; Indels 60; Gaps 12;

QY 159 LPDQWERRTASGRIQYLNHITTTOKERTRPASEYSPG-----RPLS 203
Db 246 LPSGWEERKDAKRTYYVNNHNTTTTTRPIMQAEADGASGATNSNNHLEIPEQIRRP 305
QY 204 C--FVDENTPIGTCGATCGQSSDRLAERRVRSQRHRYNMSRTHLHTPPDLPEGYEQ 260
Db 306 LSSPTVTLSPLEGAADSPVRAVKDTLSNPQSPQSPYNSPKPKVQTSFLPPGWE 365
QY 261 TTQCGQVYFHTGTGTVSTWHDPRVP-----RDLNSINCEBELGFLPPGWEIRNTATGRVYF 315
Db 366 IAFNGRPFPTIDHNTKTTTWEDPRLKFPVHWRSKTSLNPNDLGLFLPPGWEIRHLDGRFTY 425
QY 316 VDHNNRTTQFTDRLSANLHLVLRNQNLKQQQQQVSLCPDDTECLTVPR--YKRDLV 373
Db 426 IDNSKITQWEDRL-----QN-----PAITGPAPVYSREFK 457
QY 374 QKILRLQELSOQPOAGHCRIEVSREEIFEESYQVWMPKPDLMK-RLMIKFRGEGL 432
Db 458 QKYDYFRKKLKPADIPNPFEMKLRHNNIFEESYRIMSVPKRPDLKARLWIEFESEKGL 517
QY 433 DYGGVAREWLYLLSHEMLNPYGLFOYSRDDIYTLQINPD SAV-NPEHLSYHFHVGRI 491
Db 518 DYGGVAREWFFLLSKEMFNFPYGLFEYSATDNYTLQINPNSGLCNEDHLSYFTFGRVAG 577
QY 492 MAVFHGYIDGGFTLPFYKQLLGKSLTLDDELNDPDLHNSLVWILENDITGVLDHTFCV 551
Db 578 LAVFHGKLDGFFIRPFYKQMLGKQITLNDMESVSEYNSLKNILENDPT-ELDLMF 636
QY 552 EHNAYGEITQHLKPKNGKSIIPVNEENKKEYVLYVNRFLRGIEAQFLALQKGNEV 611
Db 637 DBENFGQTYQVDLKPNGSEIMVTNENKREYIDLVIQWRFVNRVQKQNAFLGFT 696
QY 612 HLLKTFDEKELELIICGLGKIDVNDKVNTRLK--HCTPDSNIVKWKAVEFFDEERRA 669
Db 697 DLIKIFDENELLCGLGDVDVNDWRQHSIYKNGIC-PNHPVIQWFWKAVLLMDAEKRI 755
QY 670 RLLQFTGSSRVPLOGKALQAGAPRLFTIHQIDACTNNLPKHAHTCFNRIDIPPE 729
Db 756 RLLQFTGTSRVPNMGFAELYGNGPQLFTIISQWGS-PEKLPRAHTCFNRDLDP 814
QY 730 KLYEKLTAIETCGF 745
Db 815 DUREKLLMAVENAQGF 830
```

Search completed: September 21, 2004, 08:04:28
Job time : 100.589 secs

This Page Blank (uspto)

Result No.	Score	Query		ID	Description
		Match	Length		
1	3939.5	97.6	735	US-08-533-205A-2	Sequence 2, Appl
2	3939.5	97.6	735	US-09-392-163A-2	Sequence 2, Appl
3	1564	38.7	766	US-08-532-205A-4	Sequence 4, Appl
4	1564	38.7	766	US-09-392-163A-4	Sequence 4, Appl
5	1354.5	33.5	927	US-08-895-601-6	Sequence 6, Appl
6	1353	33.5	834	US-08-532-205A-6	Sequence 6, Appl
7	1353	33.5	834	US-09-392-163A-6	Sequence 6, Appl
8	1336	33.1	854	US-09-070-060-4	Sequence 4, Appl
9	1336	33.1	854	US-09-357-746-4	Sequence 4, Appl
10	1329	32.9	852	US-09-070-060-3	Sequence 3, Appl
11	1329	32.9	852	US-09-357-746-3	Sequence 3, Appl
12	1298.5	32.2	906	US-08-630-916A-48	Sequence 48, Appl
13	1134.5	28.1	683	US-08-630-916A-46	Sequence 46, Appl
14	534.5	13.2	874	US-08-247-904B-8	Sequence 8, Appl
15	534.5	13.2	874	US-08-767-942A-21	Sequence 21, Appl
16	513	12.7	866	US-08-100-692-1	Sequence 1, Appl
17	513	12.7	866	US-08-674-030-1	Sequence 1, Appl
18	450.5	11.2	1083	US-08-895-601-5	Sequence 5, Appl
19	219	5.4	486	US-08-348-518C-2	Sequence 2, Appl
20	207	5.1	472	US-08-348-518C-5	Sequence 5, Appl
21	207	5.1	472	US-08-476-509B-5	Sequence 5, Appl
22	162.5	4.0	54	US-08-630-916A-124	Sequence 124, App
23	158	3.9	448	US-08-476-509B-2	Sequence 2, Appl
24	156.5	3.9	55	US-08-630-916A-75	Sequence 75, Appl
25	154.5	3.8	454	US-08-348-518C-4	Sequence 4, Appl
26	154.5	3.8	454	US-08-476-509B-4	Sequence 4, Appl
27	153	3.8	234	US-08-630-916A-50	Sequence 50, Appl

Db 48 WNOHYDLYIGKSDSVTISVNNHKK:HKQAGFLGCVRLLSNAINRLKDTGYQRLDLCKL 107
QY 121 GPNNDNTVAGQIVVLSQSDRIGTGQGVVDCSRLFDNDLPDGWEERTASGRIOYLNHIT 180
Db 108 GPNNDNTVAGQIVVLSQSDRIGTGQGVVDCSRLFDNDLPDGWEERTASGRIOYLNHIT 167
QY 181 RTQWERPRPASEYSSPGRLSCFVDNTPISGTNGATCGSSDPRLAERRVRSQRHN 240
Db 168 RTQWERPRPASEYSSPGRLSCFVDNTPISGTNGATCGSSDPRLAERRVRSQRHN 227
QY 241 YMSRTHLHTPPDLPEGVEORTTQGGVYFLHTQTGVSTWHDPRVPRDLNSINCELGPLP 300
Db 228 YMSRTHLHTPPDLPEGVEORTTQGGVYFLHTQTGVSTWHDPRVPRDLNSINCELGPLP 287
QY 301 PGWEIRNTATGRVYFVDHNNRTTQFTDPRLSANLHLVLRNQNOLKQOQOQVSLCPDDT 360
Db 288 PGWEIRNTATGRVYFVDHNNRTTQFTDPRLSANLHLVLRNQNOLKQOQOQVSLCPDDT 347
QY 361 ECLTVPRYKRDVQKILRQELSOQOQAGHCRIEVSREEIFBESYRQVMKRPKDLWK 420
Db 348 ECLTVPRYKRDVQKILRQELSOQOQAGHCRIEVSREEIFBESYRQVMKRPKDLWK 407
QY 421 RLMKFRGEGLDYGGVAREWLYLLSHEMLNPYGLFOYSRDDIYTLQINPDSAVNPEHL 480
Db 408 RLMKFRGEGLDYGGVAREWLYLLSHEMLNPYGLFOYSRDDIYTLQINPDSAVNPEHL 467
QY 481 SYFHFVGRIMGAVFHGHYIDGGFTLPFYKQLLGSITLDDMELVDPDLHNSLVWILEND 540
Db 468 SYFHFVGRIMGAVFHGHYIDGGFTLPFYKQLLGSITLDDMELVDPDLHNSLVWILEND 527
QY 541 ITGVLDHTFCVEHNAYGEIIOHELKPNKGSIPVNEENKKEYVRLYVNNRFLRGIEAQFLA 600
Db 528 ITGVLDHTFCVEHNAYGEIIOHELKPNKGSIPVNEENKKEYVRLYVNNRFLRGIEAQFLA 587
QY 601 LQKGFNEVYPOHLLKTFDEKELELIIICGLGKIDVNDKVNTRLKHCTPDSNIVKWFKAV 660
Db 588 LQKGFNEVYPOHLLKTFDEKELELIIICGLGKIDVNDKVNTRLKHCTPDSNIVKWFKAV 647
QY 661 EPPDEERRARLQFVTGSSRPVLOGFKALQAGAPRLFTIHOIDACTNNLPKHAHTCFNRI 720
Db 648 EPPDEERRARLQFVTGSSRPVLOGFKALQAGAPRLFTIHOIDACTNNLPKHAHTCFNRI 707
QY 721 DIPPYESYKLYEKLTAIETTCGFAVE 748
Db 708 DIPPYESYKLYEKLTAIETTCGFAVE 735

RESULT 2

US-09-392-163A-2
Sequence 2, Application US/09392163A
Patent No. 6503742
GENERAL INFORMATION:
APPLICANT: Beach, David H.
APPLICANT: Caliguri, Maureen
APPLICANT: Netsky, Bradley
TITLE OF INVENTION: Ubiquitin Ligases, and Uses Related Thereto
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: FOLEY, HOAG & ELIOT LLP
STREET: One Post Office Square
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109-2170
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09392,163A
FILING DATE:
PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/539,205
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Vincent, Matthew P.
REGISTRATION NUMBER: 36,709
REFERENCE/DOCKET NUMBER: CSV-005.01
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 832-1000
TELEFAX: (617) 832-7000
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 735 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-392-163A-2

Query Match 97.6%; Score 3939.5; DB 4; Length 735;
Best Local Similarity 98.0%; Pred. No. 0;
Matches 733; Conservative 0; Mismatches 2; Indels 13; Gaps 1;
QY 1 MSNPGRRNGGVKRLRLTVLCAKNLVKDDFFRLPDPFAKVVVDGSGGCHSTDTYKNTLDPK 60
Db 1 MSNPGRRNGGVKRLRLTVLCAKNLVKDDFFRLPDPFAKVVVDGSGGCHSTDTYKNTLDPK 47
QY 61 WNOHYDLYIGKSDSVTISVNNHKKIHKQAGFLGCVRLLSNAINRLKDTGYQRLDLCKL 120
Db 48 WNOHYDLYIGKSDSVTISVNNHKKIHKQAGFLGCVRLLSNAINRLKDTGYQRLDLCKL 107
QY 121 GPNNDNTVAGQIVVLSQSDRIGTGQGVVDCSRLFDNDLPDGWEERTASGRIOYLNHIT 180
Db 108 GPNNDNTVAGQIVVLSQSDRIGTGQGVVDCSRLFDNDLPDGWEERTASGRIOYLNHIT 167
QY 181 RTQWERPRPASEYSSPGRLSCFVDNTPISGTNGATCGSSDPRLAERRVRSQRHN 240
Db 168 RTQWERPRPASEYSSPGRLSCFVDNTPISGTNGATCGSSDPRLAERRVRSQRHN 227
QY 241 YMSRTHLHTPPDLPEGVEORTTQGGVYFLHTQTGVSTWHDPRVPRDLNSINCELGPLP 300
Db 228 YMSRTHLHTPPDLPEGVEORTTQGGVYFLHTQTGVSTWHDPRVPRDLNSINCELGPLP 287
QY 301 PGWEIRNTATGRVYFVDHNNRTTQFTDPRLSANLHLVLRNQNOLKQOQOQVSLCPDDT 360
Db 288 PGWEIRNTATGRVYFVDHNNRTTQFTDPRLSANLHLVLRNQNOLKQOQOQVSLCPDDT 347
QY 361 ECLTVPRYKRDVQKILRQELSOQOQAGHCRIEVSREEIFBESYRQVMKRPKDLWK 420
Db 348 ECLTVPRYKRDVQKILRQELSOQOQAGHCRIEVSREEIFBESYRQVMKRPKDLWK 407
QY 421 RLMKFRGEGLDYGGVAREWLYLLSHEMLNPYGLFOYSRDDIYTLQINPDSAVNPEHL 480
Db 408 RLMKFRGEGLDYGGVAREWLYLLSHEMLNPYGLFOYSRDDIYTLQINPDSAVNPEHL 467
QY 481 SYFHFVGRIMGAVFHGHYIDGGFTLPFYKQLLGSITLDDMELVDPDLHNSLVWILEND 540
Db 468 SYFHFVGRIMGAVFHGHYIDGGFTLPFYKQLLGSITLDDMELVDPDLHNSLVWILEND 527
QY 541 ITGVLDHTFCVEHNAYGEIIOHELKPNKGSIPVNEENKKEYVRLYVNNRFLRGIEAQFLA 600
Db 528 ITGVLDHTFCVEHNAYGEIIOHELKPNKGSIPVNEENKKEYVRLYVNNRFLRGIEAQFLA 587
QY 601 LQKGFNEVYPOHLLKTFDEKELELIIICGLGKIDVNDKVNTRLKHCTPDSNIVKWFKAV 660
Db 588 LQKGFNEVYPOHLLKTFDEKELELIIICGLGKIDVNDKVNTRLKHCTPDSNIVKWFKAV 647
QY 661 EPPDEERRARLQFVTGSSRPVLOGFKALQAGAPRLFTIHOIDACTNNLPKHAHTCFNRI 720
Db 648 EPPDEERRARLQFVTGSSRPVLOGFKALQAGAPRLFTIHOIDACTNNLPKHAHTCFNRI 707
QY 721 DIPPYESYKLYEKLTAIETTCGFAVE 748
Db 708 DIPPYESYKLYEKLTAIETTCGFAVE 735

```

RESULT 3
US-08-539-205A-4
; Sequence 4, Application US/08539205A
; Patent No. 6001619
; GENERAL INFORMATION:
; APPLICANT: Beach, David H.
; APPLICANT: Caligiuri, Maureen
; APPLICANT: Nefsky, Bradley
; TITLE OF INVENTION: Ubiquitin Ligases, and Uses Related Thereto
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FOLEY, HOAG & ELIOT LLP
; STREET: One Post Office Square
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109-2170
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/539,205A
; FILING DATE: 04-OCT-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Vincent, Matthew P.
; REGISTRATION NUMBER: 36,709
; REFERENCE/DOCKET NUMBER: CSV-005.01
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 832-1000
; TELEFAX: (617) 832-7000
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 766 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-539-205A-4

Query Match          38.7%; Score 1564; DB 3; Length 766;
Best Local Similarity 41.1%; Pred. No. 4, 1e-138;
Matches 338; Conservative 121; Mismatches 232; Indels 132; Gaps 18;

QY 1 MNPGRNRNGPKRLTVLCARMLVKKDFRLLPDPFAKVVVDGSGQCHSTDTVRNTLDPK 60
DB 1 MNSAQSR-----RIRTVIADGLYKRDYFRPDPFAVLTVGGE-QTHTTAIXKTLNPK 55
QY 61 WNOHYDLYGKSDSVTISWNHKKHKKGAGFLGCVRLLSNAINRLKDTGVQRL--DLC 118
DB 56 WNETEVAVTNDSTAIQVFDQKF-KKKGQFLGVINLRVDVLDLAIGDDEMLTRDLK 114
QY 119 KLGNPDNDTVRQIVVSLQ----- 137
DB 115 K--SNENTVHGKIILNSTTAQSTLQVPSSAASCARTQRTSITNDPQSSKSSSVSRNPA 172
QY 138 -----SDRIGTGVQVDCSLFDN-----DLPDCEWERTASGRIOVLNHLTRTQ 184
DB 173 SRAGSPTRKDNAPASPASSEPRTSFSDQYGRLLPGWERTDNLGRYTVVDHNTSTT 232
QY 185 WERTP-----RPASEYSSPCRPSCFVDENTPIGTNGATCGQSSDPLAERRVRSQRH 238
DB 233 WIRPNLSSVAGAAAEALHSSAS--SANVTEGVQPSSSNAA-----RRTEASVLT--- 279
QY 239 RNYMRTHLHPDPLPEGEYQRTTQCGQYVFLHTGTGVSTWHDPRVPRDLNIN----- 292
DB 280 ----SNATTAGSGELPPGWEQRYTPEGRPFYVDHNTTRTTTWDPRRQYIRSYGPNNA 335
QY 293 -----CEELGPLPPCWEIRNTATGRVYVDHNNRTQFTDPRLSANLHLVLRNQLKQD 347
DB 336 IQQPVSQLPLPSGWENRLNTARVYVDHNTKTTTWDDEPLPSSL-----DQ 384

```

Query Match 38.7%; Score 1564; DB 4; Length 766;

[illegible]

RESULT 5	
US-08-895-601-6	
; Sequence 6, Application US/08895601	
; Patent No. 6060262	
; GENERAL INFORMATION:	
; APPLICANT: Beer-Romero, Peggy	
; APPLICANT: Strack, Peter J.	
; APPLICANT: Glass, Susan J.	
; APPLICANT: Rolfe, Mark	
; TITLE OF INVENTION: REGULATION OF KAPPA B (1KB) DEGRADATION,	
; TITLE OF INVENTION: AND METHODS AND REAGENTS RELATED THERETO	
; NUMBER OF SEQUENCES: 16	
; CORRESPONDENCE ADDRESS:	
349: HSSRGRSLQAYTFEEQPTLPVLLPTSSGLPPGWEEKQDGRGRSYVVDNSRTTWTWKFTV 408	Qy
253: -----1PGYEORITTCQGVYELH 271	Qy
409: QATVETSQLTSSQSSAGPQASQSDSQGVQTQPSEIEQGFLPKGEVRHAPNGRPPFID 468	Db
272: TQGVSTWHDPR--VPR-----DLSNINCEELGPLPPGWEIRNTATGVRVYFVDHNKR 321	Qy
469: HNTKTTTWEDPRLKIPAHLRGKTSLDTSN-----DGLPLPPGWEERTHTDGRIFYNNHNK 524	Db
322: TTQFTDPRLSANLHLVNLNQNLKQQQQQVSVLCPDDTECLTVP--YKRDQVQKLIK 379	Qy
525: RTQWEDPRL-----ENVAITGPVPSYSDYKEKYEFF 556	Db

QY 380 ROELSCQOPQAGHCRLEVSREIFESYQVNMKRPKDLWK-RLMIKRGEGGLDYGGVA 438
 Db 557 RRLKXKQNDIPNKFEMKLRARVLEDSYRIRINGVRADFLKARLWIBFDGKGLDYGGVA 616
 QY 439 REWLILSHMLNPPYGLFQYGRDDIYTLQINPDSAV-NPEHLSYFHFVGRIMNAVPHG 497
 Db 617 REWFLISXEMFNPPYGLFEYSATDNYTLQINPNSGLCNEDELSEYKFGIRVAGNAVYHG 676
 QY 498 HYIDGGFTLPFYKOLLGKSITLDDMELVDPDLHNSVLWILENDITGVLDHFTFCVEHNAYG 557
 Db 677 KLLDGGFFIRPFYKMLHKTITLHDSVSDSEYNSLRWILENDPT-ELDLRFIDEEBFG 735
 QY 558 EIIQHELKNGKSIIPVNEENKKEYVRLYVNMWFLRGIEAQFALQKGFNEVIPPQHLKTF 617
 Db 736 QTHQHELKNGGSEIVVTKNKKKEYILYVQWRFVNRIOKQMAAFKEGFFELIPQDLIKIF 795
 QY 618 DEKELELLICGLKIDVNDWKVNRILKH-CTPDSNIVKWFKAVFFDEERARRLLQVFT 676
 Db 796 DENELELLMCGLGVDVNDWREHTYKNGYSANHQVIOQWFKAVLMDSEKRIIRLLQVFT 855
 QY 677 GSRVPLQKALQGAAGPRLTIHQIDACTNNLKAHTCFNRIDIPPVESYEKLYEKL 736
 Db 856 GTSRVPNGFAELYSNGPQSTVEQW-GTPEKLPRATCFNRDLDPPIESFEELWDLKQ 914
 QY 737 TAIECTGF 745
 Db 915 MAIENTQGF 923

RESULT 6

US-08-539-205A-6
 ; Sequence 6, Application US/08539205A
 ; Patent No. 6001619
 ; GENERAL INFORMATION:
 ; APPLICANT: Beach, David H.
 ; APPLICANT: Caligiuri, Maureen
 ; APPLICANT: Nefsky, Bradley
 ; TITLE OF INVENTION: Ubiquitin Ligases, and Uses Related Thereto
 ; NUMBER OF SEQUENCES: 6
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: FOLEY, HOAG & ELIOT LLP
 ; STREET: One Post Office Square
 ; CITY: Boston
 ; STATE: MA
 ; COUNTRY: USA
 ; ZIP: 02109-2170

COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent In Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/539,205A
 ; FILING DATE: 04-OCT-1995
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Vincent, Matthew P.
 ; REGISTRATION NUMBER: 36,709
 ; REFERENCE/DOCKET NUMBER: CSV-005.01
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (617) 832-1000
 ; TELEFAX: (617) 832-7000
 ; INFORMATION FOR SEQ ID NO: 6:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 834 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; US-08-539-205A-6

Query Match 33.5%; Score 1353; DB 3; Length 834;
 Best Local Similarity 44.6%; Pred. No. 3.5e-118;
 Matches 275; Conservative 93; Mismatches 188; Indels 60; Gaps 12;

QY 159 LPOGWERRTASGRIQYLNHITRTTOWERTRTPASEYSSPG-----RPLS 203
 Db 246 LPSGWEERKAKGRYYVYNNHNRITTTWTRFIMO.AEDGASGSATNSNNHIEPQIRPERS 305
 QY 204 C---FVDENTPISGTNGATCGSQSDPRLAERRVRSQRHRYMGRTHLHTPPDLPEGYEOR 260
 Db 306 LSSFTVLSAPLEGAKSDPVRRRAVKDTLSNPQSPQSPYNSPKPQHKVQTSFLPPGWEHR 365
 QY 261 TTQGGQVYFLHTQTGVSTWHDPRVP-----RDLSNINCEELGLPLPGWEIRNTATGRVYF 315
 Db 366 IAPNGREFFFDHNTKTTTWEDPRLKFPVHMRSKTSLNPNDLGLPLPGWEERIHLDRFTY 425
 QY 316 VDHNRRTTQTDPRLSANLHLVLRNQNLKQOQQQVSVLCPDDTECLTVPR--YKRDLV 373
 Db 426 IDHNSKITQWEDPRL-----QN-----PAITGPANVPYSREPK 457
 QY 374 QUKILQELSQOQPOQAGHCRIEVSRBEEIPEESYRQVMKMRPKDLWK-RLMIKRGEEGL 432
 Db 458 QKYDYFRKLLKPADIPNRFEMKLRHNNIPEESYRIRMSVKRDPVLKARLWIBFSEKGL 517
 QY 433 DYGGVAREWELYLASHMLNPPYGLFOYSRDDIYTLQINPDSAV-NPEHLSYFHFVGRIMG 491
 Db 518 DYGGVAREWELFLSKENFNFNYYGLFEYSATDNTLQINPNSGLCNEDELHJYFTFIGRVAG 577
 QY 492 MAVFHGYIDGGFTLPFYKOLLGKSITLDDMELVDPDLHNSVLWILENDITGVLDHFTFCV 551
 Db 578 LAVFHGKLLDGGFFIRPFYKMLGKQITLNDMESVDSYNSLWILENDPT-ELDLMFCI 636
 QY 552 EHNAYGEIIQHELKPNKSIIPVNEENKKEYVRLYVNMWFLRGIEAQFALQKGFNEVIPPQ 611
 Db 637 DEENFGQYQVLDKPNGSEIMVTNENKREYIDIVIQWRFVNRVCKQNNAFLEGTELLPI 696
 QY 612 HLLKTFDEKELELIICGLGKIDVNDWKVNRILKH-CTPDSNIVKWFKAVFFDEERARR 669
 Db 697 DLKIFDENELELLMCGLGVDVNDWRFVNRHSHIYKNGYC-PNHPVIOQWFKAVLLDAEKRI 755
 QY 670 RLLOFTVGTSSRVPLOQKALQGAAGPRLTIHQIDACTNNLKAHTCFNRIDIPPVESYE 729
 Db 756 RLLOFTVGTSSRVPNGFAELYSNGPQLFTIEQWGS-PEKLPRATCFNRDLDPPIETETE 814
 QY 730 KLYEKLTAIECTGF 745
 Db 815 DLREKLLMAVENAQGF 830

RESULT 7

US-09-392-163A-6
 ; Sequence 6, Application US/09392163A
 ; Patent No. 6503742
 ; GENERAL INFORMATION:
 ; APPLICANT: Beach, David H.
 ; APPLICANT: Caligiuri, Maureen
 ; APPLICANT: Nefsky, Bradley
 ; TITLE OF INVENTION: Ubiquitin Ligases, and Uses Related Thereto
 ; NUMBER OF SEQUENCES: 6
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: FOLEY, HOAG & ELIOT LLP
 ; STREET: One Post Office Square
 ; CITY: Boston
 ; STATE: MA
 ; COUNTRY: USA
 ; ZIP: 02109-2170
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent In Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/392,163A
 ; FILING DATE:
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 08/539,205

```

; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Vincent, Matthew P.
; REGISTRATION NUMBER: 36,709
; REFERENCE/DOCKET NUMBER: CSV-005.01
; TELEPHONE: (617) 832-1000
; TELEFAX: (617) 832-7000
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 834 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-392-163A-6

Query Match 33.5%; Score 1353; DB 4; Length 834;
Best Local Similarity 44.6%; Pred. No. 3.5e-118;
Matches 275; Conservative 93; Mismatches 188; Indels 60; Gaps 12;

QY 159 LPDGEERTASGRIOYLNHITRTQWERPTRPASEYSSPG-----RPLS 203
Db 246 LPSGWEERKDAKGRYYVYVNNHRTTTRDINQLAEDGASGATSNHNLIEPIQRPRS 305
QY 204 C---FVDENTPISTNGATCGSSDPRLAERRVRSQRHRYMSRTHLTTPDPEGYEOR 260
Db 306 LSSPTVTLSAPLEGAKDSPVRRRAVKOTLSNPQSPQSPYNSPKQHKVQTQSFLPPGWEWR 365
QY 261 TTQGGQVYFHTQTGVSTWHDPRVP-----RD-SNINCEBELGPPGWEIRNTATGRVYF 315
Db 366 IAPNGRFFFDHNTKTTTWEDPRLLKPPVHMSKTSLNPNLDGLPLPGWEERIEHLDKRTFY 425
QY 316 VDHNRRTQFTDPLSANHLHLVLRNQKLDKQOQOQVSLCPDDTECLTVPR--YKRDLV 373
Db 426 IDHNSKITOWEDPRL-----QN-----PAITGFAPVPSREFK 457
QY 374 QKLKILRQELSOQOQAGHCKRIEVSREIPEESYRQVMKVRPKDLMK-RLMIKFRGEEGL 432
Db 458 QKYDYFRKKLKKPADIPNRFEMKLRNNIPEESYRRIMS VKRDPVLKARLWIFSESEKGL 517
QY 433 DYGVAREWLYLLSHEMLNPYGLFQYSRDDITYLQINPDSAV-NPEHLSYFHFVRIMG 491
Db 518 DYGVAREWFLLSKEMFNPFYGLFYSATDNTYLQINPNSGLCNEHDLSYFTFGRVAG 577
QY 492 MAVFHGHYDGGFTLPFKQLLGSITLDDMELVDPDLNLSLVILENDITGLDHTFCV 551
Db 578 LAVFHGKLDDGFTLRPFYKMLGKQIITLNDVESVDSYNSLKWILENDPT-ELDLMFCI 636
QY 552 EHNAYGEIIQHELKPKNGKSI PVNEENKEVYRVLVYNNRFLRGIEAQFLALQKGFNEVPIQ 611
Db 637 DEENFGQTYQVDLPKNGSEIMVTNENKREYIDLVIQWRFNRYQKQMNAFLEGTELLPI 696
QY 612 HLLKTFDEKELELLICGLKIDNDKVNTRLK--HCTPDSNIVKFWKAVESFFDEERRA 669
Db 697 DLKIFDENELELLMCGLDGVDVNDWRQHSYKNGYC-PNRPVQWFKAVLLMDAEKRI 755
QY 670 RLLQFVTGSSRVPLQGGKALQAGAPRLFIHQIDACTNNLPKHAHTCFNRIDIPPYESYE 729
Db 756 RLLQFVTGTSRVPNNGFAELVGSNGPQLFTIEQWS-PEKLPRAHTCFNRLDLPYETPE 814
QY 730 KLYEKLTAIETCGF 745
Db 815 DLREKLLMAVENAQGF 830

RESULT 8
US-09-070-060-4
; Sequence 4, Application US/09070060
; Patent No. 5976849
; GENERAL INFORMATION:
; APPLICANT: Husted, Carolyn M.
; APPLICANT: Ghildyal, Namit
; TITLE OF INVENTION: Human E3 Ubiquitin Protein

```

```

; TITLE OF INVENTION: Ligase
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ZENECA Pharmaceuticals, Inc.
; STREET: 1800 Concord Pike
; CITY: Wilmington
; STATE: DE
; COUNTRY: USA
; ZIP: 19850-5437
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/070,060
; FILING DATE: 30-APR-1998
; CLASSIFICATION:
; PRIOR APPLICATION NUMBER: 60/073,839
; FILING DATE: 05-FEB-1998
; ATTORNEY/AGENT INFORMATION:
; NAME: Higgins, Patrick H
; REGISTRATION NUMBER: 39,709
; REFERENCE/DOCKET NUMBER: PHM.70312
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 302.886.4889
; TELEFAX: 302.886.8221
; TELEX:
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 854 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: peptide
; US-09-070-060-4

Query Match 33.1%; Score 1336; DB 2; Length 854;
Best Local Similarity 36.6%; Pred. No. 1.4e-116;
Matches 316; Conservative 130; Mismatches 272; Indels 146; Gaps 24;

QY 13 KURLTVLCAK-NLVKKDPFRLPDPFAKVVVGSGCQCHSTVKTLPKKNQHYDLVIGK 71
Db 9 QLIQTVISAKLKKKKWFG-PSPYVEVTVD--QSKKTEKNNNSPKMKQPLTVIVTP 65
QY 72 SDSVTISVYNNHKKIKK--QGAGFLGCVRLLSNAINRLKDTGYQRLDLCKLGPNDNDTVR 129
Db 66 TSKLCFRVWSHQTLKSDVLLGTAGLDIYETLKSNNMKLEEV---VMTLQLVGDKEPTETM 122
QY 130 GQIVY---SLQSRDRIGTGGQVY-----DCSRLFDNLDP-----GWEE 165
Db 123 GDLSVCLDGLQVAEAVTNGETSCSESTQNDGCRTRDDTRVSTNGSEDEPEVAASGENK 182
QY 166 RTASGRQYLNHITRTTQWERPTRPASEYSSGRLPSCFVDENT-----PIS 213
Db 183 RANGNNSPSLNGGFKSPRPPRPPPP--PTPRPASVNGSPSTNSDSGSSSTGSLPPT 240
QY 214 GTN-----GATCG-----QSSDPLAERRVRSQRHNYM-----242
Db 241 NTNVNTS:SEGATSLIIPLTISGSGSPRLNTVSQAPLPGEQRV-DQGRVYVVDHV 299
QY 243 -SRTHLHTPPDLPEGYEORTTQCGOVYELHTQTGVSTWHDPRVP-----285
Db 300 EXRTTDRPEPLPGWEERRVDNMGRIYYVDHFTTITTTQWRPTLSSVRNYEQWQLQRSLQ 359
QY 286 -----RDL-----SNINCEBELGPPGWEIRNTATGRVYFVDHNNRTQFTD 327
Db 360 GAWQFNQRFYVGNQDLFATSONKEFDPLGPLPGWEKRTDSNGRVYFVNHTITQWED 419
QY 328 PRLSANHLVLRNQKLDKQOQOQVVSLLCP---DDTECLTV---PR-----Y 368
Db 420 PRSQGQ-----LNEKPLPEGWEWRFTVDGIPYFVDHNNRATTYIDPRTOKSALDNGPQIAY 475

```


Db 458 PRTGKS---ALONGPOI-----AYVRDPKAKVQYFRFWCQOLA 492
 QY 385 OQPOAGHCRIEVSREIEFEESYOVWKMMPKDLWKRLMIKFRGEGLDYGVAREWYVL 444
 Db 493 MPQ-----HIKIVTRKTLFEDSFQIMSPSQDLRRLLWVIFPGEGLDYGVAREWYVL 548
 QY 445 LSHMLNPPYGLFQYSRDDIYIQLNPDSDAVNPEHLSYFHVGRIMGMAVFGHYIDGGF 504
 Db 549 LSHVNLNPPYCLFVAGKDNVYCLQINPASYINPDHLKVPFRFICRTIAMLPHFKGIDTGF 608
 QY 505 TLFPYKOLGKGLITLDDMLVDPDLHNSLVWTLNDITGV-LDHTFCVHEHNAAYGIIIOHE 563
 Db 609 SUPFYKRLINKVGLXDLJESIDPEFVNSLIWKENNIECDLEWYFSDVKELGIIKSHD 668
 QY 564 LKPNKSIPIVNEENKEYVRLVYNNRFLGIAQFLALQKGFNEVPIQELKLTDFDEKELE 623
 Db 669 LKPNNGNIIYVTEENKEEYIRMAEVRLSRGVEEQTAQAFPEGNEIPLQOYLQYFPAKELE 728
 QY 624 LIICGLGKIDVNDKVNTRKCTPDSNVKFWKAVEFFDEERBARLLQFVTGSSRVL 683
 Db 729 VLICGMEIDLDNDQWHAIRYVARTSKQIMWFQFVKEIDNEKRWLLQFVTGTCRLPV 788
 QY 684 QGFKALQAGAPRLFTIHQIDACTNNLPKHAHTCFNRIDIPPYESYKLYEKLITAEIETC 743
 Db 789 GGFADLMGSGKQFCIEKVGK-ENWLPKSHCTFNRDLDPYKSYEQLEKLLPAIEETE 847
 QY 744 GRAVE 748
 Db 848 GFGQE 852

RESULT 12

US-08-630-916A-48
 ; Sequence 48, Application US/08630916A
 ; Patent No. 601137
 ; GENERAL INFORMATION:
 ; APPLICANT: Pirozzi, Gregorio
 ; APPLICANT: Kay, Brian K.
 ; APPLICANT: Fowlkes, Dana M.
 ; TITLE OF INVENTION: IDENTIFICATION AND ISOLATION OF NOVEL
 ; TITLE OF INVENTION: POLYPEPTIDES HAVING WW DOMAINS AND METHODS OF USING SAME
 ; NUMBER OF SEQUENCES: 124
 ; CORRESPONDENCE ADDRESS:
 ; STREET: 1155 Avenue of the Americas
 ; CITY: New York
 ; STATE: New York
 ; COUNTRY: United States
 ; ZIP: 10036-2711
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/630,916A
 ; FILING DATE: 03-APR-1996
 ; CLASSIFICATION: 435
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: MISROCK, S. LESLIE
 ; REGISTRATION NUMBER: 18,872
 ; REFERENCE/DOCKET NUMBER: 1101-203
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (212) 790-9090
 ; TELEFAX: (212) 896-8864/9741
 ; INFORMATION FOR SEQ ID NO: 48:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 906 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS:
 ; TOPOLOGY: unknown
 ; MOLECULE TYPE: peptide
 ; US-08-630-916A-48

Query Match 32.2%; Score 1298.5; DS 3; Length 906;
 Best Local Similarity 42.0%; Pred. No. 5.4e-113;
 Matches 271; Conservative 101; Mismatches 216; Indels 57; Gaps 9;

QY 122 PNDNDTVRGQIVVLSQSRDRIGTGGVVDVCSRLFDNDLPDGMBERRTASGRIOYLNHTR 181
 Db 301 PNPNTTSLPAPATPAEGEPESTSGTQOLPAAQAPDALPAGWEGRELNGRVVYVDHNTK 360
 QY 182 TTQWERPTRASEYSSPGRPLSCFVDENTISGTINGATCGSSDPLRAERRVRSORHN- 240
 Db 361 TTTWERPLPFGWEKRTDPRGRFYVDHNT-----RTTWQRPTAEYVRVYEQWSORNQ 414
 QY 241 -----YMSRTHLHTP-PDLPEGYEQRTOQGOVYFHLTGTGVSTWHDPRV 284
 Db 415 LQGAHQFSQRFLYQFWSASTDHLPLPFGWEKR-QDNGRVVYVNHNTTQTOWEDPRT 473
 QY 285 PRDLSNINCEBELGFLPGLWEIRNATGRVYVDHNRRTQFTDPRLSANLEHLVLRNQQL 344
 Db 474 QGMIOE-----PALPFGWEMKVTSEGVRYVDHNTTTTTFKDPRPGE----- 516
 QY 345 XDQOQOQVWSLCPDDTECLTVPRYKRDVLQKILRQELSQQQFQAGHCRIEVSREIFE 404
 Db 517 -----SGTKQSGPGAYDRSFRWKYHQFR-FLCHSNALPSHVKISVSRQLTPE 562
 QY 405 ESYROWWMPKDLWKRLMIKFRGEGLDYGVAREWYLLSHMLNPPYGLTQYSRDDI 464
 Db 563 DSFOQIMMKPYDLRRLYIIMRGEGLDYGGIAREWFFLLSHEVLNPPYCLFEPYAKNN 622
 QY 465 YTLQINPDSAVNPEHLSYFHVGRIMGMAVFGHYIDGGFTLFPYKOLGKSLITLDDMEL 524
 Db 623 YCLOINPASSINPDHLTYFRFICRTIAMLPHFKGIDTGTFTLFPYKMLNKRPTLKDL 682
 QY 525 VDPDLHNSLVILENDITGV-LDHTFCVHEHNAAYGIIQHELKPNKGSIPVNEENKKEYVR 583
 Db 683 IDPEFYNSIWMKENNLECGLELYFTQDMEILLKGVTHLEKEGSEIRVTEENKEEYIM 742
 QY 584 LYNNWRELGRIEAQFLALQKGFNEVPIQHLKLTDFDEKELELIICGLGKIDVNDKVNTRL 643
 Db 743 LLTDWRTRTGVESQTKAFLDGFNEVAPLEWIRYFDEKELEMLCGLMCGMEIDMSQKSTIY 802
 QY 644 KHCTPDSNVKFWKAVEFFDEERBARLLQFVTGSSRVPVLPQGFALQAGAPRLFTIHOI 703
 Db 803 RHVTKNSKQIQWFWQVVKEMDNKRIQLQFVTGTCRLPVGGAELIGSNGPKQFCIDKV 862
 QY 704 DACTNNLPKHAHTCFNRIDIPPYESYKLYEKLITAEIETCFAVE 748
 Db 863 GKET-WLPRSHTCFNRDLDPYKSYEQLEKLLYAEIETEFGQGE 906

RESULT 13

US-08-630-916A-46
 ; Sequence 46, Application US/08630916A
 ; Patent No. 601137
 ; GENERAL INFORMATION:
 ; APPLICANT: Pirozzi, Gregorio
 ; APPLICANT: Kay, Brian K.
 ; APPLICANT: Fowlkes, Dana M.
 ; TITLE OF INVENTION: IDENTIFICATION AND ISOLATION OF NOVEL
 ; TITLE OF INVENTION: POLYPEPTIDES HAVING WW DOMAINS AND METHODS OF USING SAME
 ; NUMBER OF SEQUENCES: 124
 ; CORRESPONDENCE ADDRESS:
 ; STREET: 1155 Avenue of the Americas
 ; CITY: New York
 ; STATE: New York
 ; COUNTRY: United States
 ; ZIP: 10036-2711
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; OPERATING SYSTEM: IBM PC compatible
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/630,916A
 ; FILING DATE: 03-APR-1996
 ; CLASSIFICATION: 435
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: MISROCK, S. LESLIE
 ; REGISTRATION NUMBER: 19,872
 ; REFERENCE/DOCKET NUMBER: 1101-203
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (212) 790-9090
 ; TELEFAX: (212) 896-8864/9741
 ; INFORMATION FOR SEQ ID NO: 46:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 683 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS:
 ; TOPOLOGY: unknown
 ; MOLECULE TYPE: peptide
 ; US-08-630-916A-46

Query Match 28.1%; Score 1134.5; DB 3; Length 683;
 Best Local Similarity 43.7%; Pred. No. 9.4e-98;
 Matches 241; Conservative 84; Mismatches 182; Indels 45; Gaps 11;
 QY 159 LPDGEERRTAGSRIQYLNHITRTQWER--PTRPASEYSSPORPLSCFVDENPTISGNT 216
 DB 162 LPSCGEQRKDPHGRTYVVDHNTTRTTTWERPOPLPGWERRVDDRRYVYVDHNTTRTTWQ 221
 QY 217 GATC-----GQSDPLPLAERRVRSQRHNYM--SRTHLHTPPD---LPEGEQRTTQQG 265
 DB 222 RPTMESVRNFEQWQSQNRQNGAQMQNQRYLYSASMLAENDPYGLPPGWEXKRVDSST 281
 QY 266 QVYELHTQTGVSTVHWDPRVPRDLNSINCEBELGPLPGCWEIRNTATGRVYFVDHNNRTTF 325
 DB 282 RVYFVNNTKTQWEDRT--QGLQN---EE--PLPEGWEIRYTRGVRYFVDHNTTRTTTF 335
 QY 326 TDPRLSANLHLVLRNQLKDDQQQQVVSICPDDTECLTVPRYKRDVLQKILKRLQELS 385
 DB 336 KDPK-----NGKSVTKGSPQIA-----YERGRWKLAFRY--LCQ 370
 QY 386 QPQAGHCRIVSREEIFESYRVQVMKMRPKDLWKRLMKFRGEEGLDYGVGAREWLYLL 445
 DB 371 SNALPSHVKNVSRQTLFESFQOIMALKYDRLRRLLYVIFRGEGLDYGLAREWFFLL 430
 QY 446 SHEMLNPYGLFQYSRDDIYTLQNPDSAVNPEHLSYFHFVGRIMGNVPHGYIDGGFT 505
 DB 431 SHEVLNPMYCLFEYAGKNNYCLQINPASTINPDHLSYFCIGRFIAVALPHGKFIDTGS 490
 QY 506 LPFYKQLIGKSITLDDMELVDPDLHNSLVMTLENDITGV--LDHTFCVEHNAVGEIIOHEL 564
 DB 491 LPFYKMLSKKLTIKDESIDTEFNSLIWTRDNNIECGLEMYFSVDMELKGVTSHDL 550
 QY 565 KPNKSIPIVNEENKEYRYLVYVNRFLRGIEAQLAKQGFNEVIPOHLLKTTDEXELEL 624
 DB 551 KLGSSNLLVTEENKDEYIGLMTWERSRGVQEQTKAFLDGFNEVWPLQWLYFDEKELEV 610
 QY 625 IICGLKIDVNDKVNTRKCHCTPDSNIVKWFKAZVFFDEERRARLLOFVTCSSRPLQ 684
 DB 611 MLCGQVEVDLADQWRVTYRHYTRNSKQIIFWFQVKTENEVRMLQLQFVTGTCRLPLG 670
 QY 685 GFALQGAAGPR 696
 DB 671 GFALMSGNGPR 682

RESULT 14

; US-08-247-904B-8
 ; Sequence 8, Application US/08247904B
 ; Patent No. 5981699
 ; GENERAL INFORMATION:
 ; APPLICANT: Rolfe, Mark
 ; APPLICANT: Eckstein, Jens W.
 ; APPLICANT: Draetta, Giulio

; TITLE OF INVENTION: Human Ubiquitin Conjugating Enzyme
 ; NUMBER OF SEQUENCES: 17
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Foley, Hoag & Eliot
 ; STREET: One Post Office Square
 ; CITY: Boston
 ; STATE: MA
 ; COUNTRY: USA
 ; ZIP: 02109
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: ASCII(text)
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/247,904B
 ; FILING DATE: 23-MAY-1994
 ; CLASSIFICATION: 530
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Vincent, Matthew P.
 ; REGISTRATION NUMBER: 36,709
 ; REFERENCE/DOCKET NUMBER: MIV-029.01
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (617) 832-1000
 ; TELEFAX: (617) 832-7000
 ; INFORMATION FOR SEQ ID NO: 8:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 874 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; US-08-247-904B-8
 Query Match 13.2%; Score 534.5; DB 2; Length 874;
 Best Local Similarity 25.1%; Pred. No. 4.5e-41;
 Matches 185; Conservative 117; Mismatches 275; Indels 161; Gaps 23;
 QY 105 NRLKDTGQRDLDCVGN---DNDTVRGQIVVSLOSRRDRTGTG-----GQVVVDCS 152
 DB 202 SRIGSSQGDNNLQKGLDGDVSDVIDAIR--RVYTELLSNEKIEAFNLALVYLSPNVECD 260
 QY 153 RLFDNDLPDGEERRTAGSRIQYLNHITRTQWERPTRPASEYSSPORPLSCFVDENPTI 212
 DB 261 LTYHN-----VYSEDENLNLF--IICMENRNLHSPEYLEMALPLFCCKAMSKLPL 308
 QY 213 SGTNG-----ATCGQSSDPLAERRVRSQRHNYMSRTHL----- 247
 DB 309 AAQKGLIBLSKYNADQIRRMWETFQQLITYKVISNBFNSRNLVNEFNSRNLVNDDDAIV 368
 QY 248 -----HTPPD-----LPS-----GYEORTTQQQVYPLH 271
 DB 369 AASKCLKMYVYANVVGGEVDINNEEDDEEPIPESSSELTQLQELLGEEERNNKGLRVDPLE 428
 QY 272 TOTGVSTVHWDPRVPRDLNSINCEBELGPLPGCWEIRNTATGRV-----YFVDHNNRT 322
 DB 429 TELGVKT-----LDCKK--PLIPFEEFINEPLNEVLENDKDYTFKVFETENKF 474
 QY 323 TOFTDPRL--SANLHLVLRNQLKDDQQQQVVSICPDDTECLTVPRYKRDVLQKILR 380
 DB 475 SMTCPFFILNAVTKNLGLYDNRIRMYSERITVL----- 509
 QY 381 QELSQQQPQAGHCRIVSREEIFES-----YQVWMOGRPKDLWKRLMKFRGEEGLDYGV 437
 DB 510 YSLVQVQQLNPYLRLKVRDHIIDALVRLMIAMENPADLKQLYVEFEQEGVDEGV 569
 QY 438 AEWLWLLSHEMLNPYGLFOYSRDDIYTLQNPDSAVNPEHLSYFHFVGRIMGNVPHG 497
 DB 570 SKEFFQLVVEEINFNDIGMFTYD--ESTKLFWNFESSF---ETEGQFTLIGVIGLAIYN 625
 QY 498 HYIDGGFTLPFYKQLLGKSIITLDDMELVDPDLHNSLVWILE--NDITGLVDHTTCV--EHN 554
 DB 626 CILDVHFPWVYVYKLMGKGLFVLDLGDGSHPVLYQSLKDLLEYVGVNVEDDMMITFIQISQT 685

QY 555 AYGEIIQHELKPKNGKSIPIVNEENKKEYVRLVYNWFLRGIEBAQFALQKGFNEVIPHQLL 614
DB 686 LFGNPMYDLKENGDKIPITNENRKEFVNLYSDIYLKNSVEKQFKAFRRGFHMVTNBSPL 745
QY 615 K-TFDEKELELIICGLKGVNDVKNVRLK-HCTPDSNIVKFWKAVFEFFDEERRALL 672
DB 746 KLFPRPEIEILLIGSRNLDFOALEBTEYDGGYTRDSVLIREFWEIVHSFTDEOKRLFL 805
QY 673 QFTGSSRVPLQG---FKALQGAAGPRLFTIHQIDACTNNLPAKHTCFNRIDIPYESYE 729
DB 806 QFTGTDRAPVGGGLKMKMIKANGPD-----TERLPTSHTCFNVLLLPYSSKE 855
QY 730 KLYEKLLTAIEETCGFAV 747
DB 856 KLERLLKAITYAKGFGM 873

RESULT 15
US-08-767-942A-21
Sequence 21, Application US/08767942A
Patent No. 6068982
GENERAL INFORMATION:
APPLICANT: Rolfe, Mark
APPLICANT: Chiu, M. Isabel
APPLICANT: Berlin, Vivian
APPLICANT: Damagnez, Veronique
APPLICANT: Draetta, Giulio
APPLICANT: Guillaume, Cottarel
TITLE OF INVENTION: UBIQUITIN CONJUGATING ENZYMES
NUMBER OF SEQUENCES: 45

CORRESPONDENCE ADDRESS:
ADDRESSEE: FOLEY, HOAG & ELIOT LLP
STREET: One Post Office Square
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109-2170
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/767,942A
FILING DATE: 17-DEC-1996
ATTORNEY/AGENT INFORMATION:
NAME: Vincent, Matthew P.
REGISTRATION NUMBER: 36,709
REFERENCE/DOCKET NUMBER: MIV-029.04
TELEPHONE: 617-832-1000
TELEFAX: 617-832-7000
INFORMATION FOR SEQ ID NO: 21:
SEQUENCE CHARACTERISTICS:
LENGTH: 874 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-767-942A-21

Query Match 13.2%; Score 534.5; DB 3; Length 874;
Best Local Similarity 25.1%; Pred. No. 4.5e-41;
Matches 195; Conservative 117; Mismatches 275; Indels 161; Gaps 23;
QY 105 NPLKDTGYORLDLCKLGNP-----DNDTVRGQIVWSLQSRDRIGTG-----GOVYDCS 152
DB 202 SRIGSSQGNLNLQKLPDSDVSDIDAIR-RVYTRLLSNEKIETAFNLALVYLSNVNVEDC 260
QY 153 RLFDNDLPDGEERRPASPGRIOYLNHITRTTOWERPTPASEYSSGPRPLSCFVDENTPI 212
DB 261 LYHN-----VYSDPNLNLFP--IIGMENRNLHSPLEYEMALPLFCCKAMSKLPL 308
QY 213 SCTNG-----ATCQSSDPLAERRVRSQFHRNMGRTHL----- 247

DB 309 AAQKGLRLMSKYNADOIRRMETFOQLITYKVIISNFBNSRNLVNEFNSRNLVNDDAIV 368
QY 248 -----HTPPD-----LPE-----GYEQRITQQQVYFLH 271
DB 369 AASKCLMWVYANVVGGEVDTNHNEEDDEBPIDESSLTI-QELLGEBERRNKKGLRVDPLE 428
QY 272 TOTGVSTWHDPRVPRDLSINCEBELGLPFGWEIRNTATGRV-----YFVDHNNRT 322
DB 429 TELGVKT-----LDCRX--ELIPFEFINEPLNEVLEMDKDYTFKFVETENKF 474
QY 323 TOFTDPLR--SANLHLVLRNQNLKDOQOOVSVSLCPDDTECLTVPRYKRDVLQKILR 380
DB 475 SFMTCFILLNAVTKNLGLYYDNIRMYSERITVL----- 509
QY 381 QELSQQOPOQACHRIEVSREIFEES---YRQVNMKMPKDLWKPLMKPKRGEELDYGCV 437
DB 510 YSLVQGGQLNPLYRLKVRDRHIIIDDALVRLEMIAMENPADLKQLYVEFEQEGQVDEGGV 569
QY 438 AREWLVLLSHEMLNPPYGLFOYSRDDIYTIQINPDSAVNPEHLSYHFHVGRIIMGMAVPHG 497
DB 570 SKEFFQLVVEEIFNPDIGMFTYD-ESTKLPWFNPSSF---ETEGQFTLIGIVGLALYNN 625
QY 498 HYIDGGFTLPFYKOLLGKSITLDDMELVDPDLNLSLVILE--NDITGVLDHTFCV-EHN 554
DB 626 CILDVHFPMVYVKLMGKGLFVDLGDSEHVLVQSLKDLLEYVGNVEDDMWITFQISQTN 685
QY 555 AYGEIIQHELKPKNGKSIPIVNEENKKEYVRLVYNWFLRGIEBAQFALQKGFNEVIPHQLL 614
DB 686 LFGNPMYDLKENGDKIPITNENRKEFVNLYSDIYLKNSVEKQFKAFRRGFHMVTNBSPL 745
QY 615 K-TFDEKELELIICGLKGVNDVKNVRLK-HCTPDSNIVKFWKAVFEFFDEERRALL 672
DB 746 KLFPRPEIEILLIGSRNLDFOALEBTEYDGGYTRDSVLIREFWEIVHSFTDEOKRLFL 805
QY 673 QFTGSSRVPLQG---FKALQGAAGPRLFTIHQIDACTNNLPAKHTCFNRIDIPYESYE 729
DB 806 QFTGTDRAPVGGGLKMKMIKANGPD-----TERLPTSHTCFNVLLLPYSSKE 855
QY 730 KLYEKLLTAIEETCGFAV 747
DB 856 KLERLLKAITYAKGFGM 873

Search completed: September 21, 2004, 07:49:22
Job time : 28.8994 secs

This Page Blank (uspo)